

1995

Quadratically constrained convex quadratic programmes.

Wiesława Teresa. Obuchowska
University of Windsor

Follow this and additional works at: <http://scholar.uwindsor.ca/etd>

Recommended Citation

Obuchowska, Wiesława Teresa., "Quadratically constrained convex quadratic programmes." (1995). *Electronic Theses and Dissertations*. Paper 3249.

This online database contains the full-text of PhD dissertations and Masters' theses of University of Windsor students from 1954 forward. These documents are made available for personal study and research purposes only, in accordance with the Canadian Copyright Act and the Creative Commons license—CC BY-NC-ND (Attribution, Non-Commercial, No Derivative Works). Under this license, works must always be attributed to the copyright holder (original author), cannot be used for any commercial purposes, and may not be altered. Any other use would require the permission of the copyright holder. Students may inquire about withdrawing their dissertation and/or thesis from this database. For additional inquiries, please contact the repository administrator via email (scholarship@uwindsor.ca) or by telephone at 519-253-3000ext. 3208.



National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

**QUADRATICALLY CONSTRAINED CONVEX
QUADRATIC PROGRAMMES**

by

Wiesława T. Obuchowska

A Dissertation

submitted to the Faculty of Graduate Studies and Research

through the Department of Mathematics and Statistics

in partial fulfillment of the requirements for the

degree of Doctor of Philosophy at

the University of Windsor

Windsor, Ontario, Canada

January, 1995



National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

THE AUTHOR HAS GRANTED AN
IRREVOCABLE NON-EXCLUSIVE
LICENCE ALLOWING THE NATIONAL
LIBRARY OF CANADA TO
REPRODUCE, LOAN, DISTRIBUTE OR
SELL COPIES OF HIS/HER THESIS BY
ANY MEANS AND IN ANY FORM OR
FORMAT, MAKING THIS THESIS
AVAILABLE TO INTERESTED
PERSONS.

L'AUTEUR A ACCORDE UNE LICENCE
IRREVOCABLE ET NON EXCLUSIVE
PERMETTANT A LA BIBLIOTHEQUE
NATIONALE DU CANADA DE
REPRODUIRE, PRETER, DISTRIBUER
OU VENDRE DES COPIES DE SA
THESE DE QUELQUE MANIERE ET
SOUS QUELQUE FORME QUE CE SOIT
POUR METTRE DES EXEMPLAIRES DE
CETTE THESE A LA DISPOSITION DES
PERSONNE INTERESSEES.

THE AUTHOR RETAINS OWNERSHIP
OF THE COPYRIGHT IN HIS/HER
THESIS. NEITHER THE THESIS NOR
SUBSTANTIAL EXTRACTS FROM IT
MAY BE PRINTED OR OTHERWISE
REPRODUCED WITHOUT HIS/HER
PERMISSION.

L'AUTEUR CONSERVE LA PROPRIETE
DU DROIT D'AUTEUR QUI PROTEGE
SA THESE. NI LA THESE NI DES
EXTRAITS SUBSTANTIELS DE CELLE-
CI NE DOIVENT ETRE IMPRIMES OU
AUTREMENT REPRODUITS SANS SON
AUTORISATION.

ISBN 0-612-01472-X

Canada

WIESŁAWA T. OBUCHOWSKA, All Rights Reserved, 1995

ABSTRACT

QUADRATICALLY CONSTRAINED CONVEX QUADRATIC PROGRAMMES

In this thesis we consider four problems arising from our study of quadratically constrained convex quadratic programmes (QCQP). The first problem concerns the representation of a quadratically constrained convex feasible region. We define the term “minimal representation” and give necessary and sufficient conditions for a representation to be minimal.

The second problem deals with boundedness of quadratically constrained convex feasible regions. In particular, we provide necessary and sufficient conditions for unboundedness which are in the form of an algorithm which requires the identification of implicit equality constraints in homogeneous linear systems.

The third problem is concerned with feasible regions which are either unbounded or not full dimensional; and with representations of these regions which may contain redundant constraints or pseudo-quadratic constraints. (A pseudo-quadratic constraint is a quadratic constraint that can be replaced with a finite number of linear inequality constraints). We show how a method of centers can be modified to solve problems with unbounded feasible regions, even in the case when no analytic center exists. We also give an algorithm to transform the problem into an equivalent problem which has a full dimensional feasible region in a lower dimensional space. Provided that an initial feasible point is given, the algorithm can also be used to find an initial interior point. Our methods of dealing with this third problem are unique in that they involve neither the addition of extra constraints nor the addition

of extra variables.

Finally we develop a new method for improving the rate of convergence of analytic center methods and of barrier function methods applied to QCQP. The underlying ideas lead to a superlinearly convergent predictor-corrector algorithm.

Dedicated to my Parents and my Family

ACKNOWLEDGEMENTS

I would like to acknowledge the help I received from many people and institutions throughout my studies.

First of all I would like to express my most sincere thanks and appreciation to my supervisor, Professor Richard J. Caron, for his valuable guidance and consideration throughout the course of this research. His help and friendship during my years here will never be forgotten.

I am also grateful to my family for their support, encouragement and understanding through all the years of my graduate program.

Many thanks to my external examiner Professor Katta Murty, not only for giving this thesis his consideration, but also for his suggestion to work with him on extensions of the topics in this thesis.

I would like to thank the members of the examining committee, Dr. J. F. McDonald, Dr. H.R. Atkinson and Dr. Y.P. Aneja for their time spent reading this thesis.

Finally, I would like to extend my thanks to Dr. H. Wolkowicz of the University of Waterloo for providing reference material related to the thesis.

The research for this thesis was supported by an NSERC Postgraduate Scholarship, an Ontario Graduate Scholarship and by various University of Windsor Scholarships. In addition, support was also received through Dr. Caron's NSERC research grant No. A8808. I am very grateful for this support.

TABLE OF CONTENTS

ABSTRACT	iv
DEDICATION	vi
ACKNOWLEDGEMENTS	vii
LIST OF TABLES	xi
LIST OF FIGURES	xi
GLOSSARY	xii
CHAPTER ONE	
INTRODUCTION	
1.1	Overview and Outline of Thesis 1
1.2	Problem Definition and Background Material 5
1.2.1	Logarithmic Barrier Function Method and Method of Centers 6
1.2.1.1	A Logarithmic Barrier Function Method: Algorithm A 7
1.2.1.2	A Method of Centers: Algorithm B 8
1.2.1.3	Remarks on Algorithms A and B 10
1.2.2	Implicit Equalities in Homogeneous Linear Systems 10
CHAPTER TWO	
MINIMAL REPRESENTATION	
2.1	Introduction 13
2.2	Implicit Equality Constraints 16
2.3	Pseudoquadratic Constraints 20
2.4	Faces of \mathcal{R} 23

2.5	Minimal Representation Theorem	29
2.6	Identification of a Minimal Representation	31
2.7	Summary	33

CHAPTER THREE

BOUNDEDNESS

3.1	Introduction	35
3.2	The Algorithm	37
3.3	Examples	42
3.4	Summary	45

CHAPTER FOUR

FAULTY FEASIBLE REGIONS

4.1	Introduction	46
4.2	Unbounded Feasible Regions	46
4.3	Feasible Regions That Are Not Full Dimensional: Initial Interior Point	54
4.4	Faulty Representations	55
4.5	Summary	58
	APPENDIX 4A	59

CHAPTER FIVE

IMPROVING THE RATE OF CONVERGENCE

5.1	Introduction	60
5.2	The Auxiliary Sequence	62
5.3	Superlinearly Convergent Algorithm	68
5.4	Computational Results	72
5.5	Summary	75

CHAPTER SIX

CONCLUDING REMARKS

6.1	Introduction	77
6.2	Contribution of Thesis	77
6.3	Future Research Directions	78

FIGURES	80
----------------	----

REFERENCES	88
-------------------	----

VITA AUCTORIS	94
----------------------	----

LIST OF TABLES

TABLE 1: Number of iterations and Newton steps

TABLE 2: The values of $\frac{c^T(x^k - x^*)}{c^T(x_0 - x^*)}$

LIST OF FIGURES

FIGURE 2.1 (a): Constraint (4) can be replaced with linear constraints.

FIGURE 2.1(b): Constraints (2) and (3) are redundant.

FIGURE 2.1(c): A minimal representation of \mathcal{R} .

FIGURE 2.1(d): A face, \mathcal{F}_1 , having more than one component.

FIGURE 2.2: The quadratic constraint can be replaced with one linear equality.

FIGURE 2.3: The quadratic constraint can be replaced with two linear inequality constraints.

FIGURE 2.4: Illustration for Algorithm C.

FIGURE 2.5: Illustration for the proof of Theorem 2.4.2

FIGURE 3.1: Illustration for Example 3.1.

FIGURE 4.1: $\mathcal{R}(z)$ has an infinity of analytic centers.

FIGURE 4.2: $\mathcal{R}(z)$ has no analytic center but the QCQP has a solution.

GLOSSARY

Symbol	Meaning
a_i	Coefficients of linear term in i-th constraint
B_i	Hessian matrix in i-th constraint
b_i	Scalar term in i-th constraint
$Q(x)$	Objective function
$Q_i(x)$	Quadratic function defining i-th constraint
S_i	Surface of i-th constraint
\mathcal{F}_i	Face corresponding to i-th constraint
$L(x, r)$	Logarithmic barrier function
$\mathcal{L}(x, u)$	Lagrangian function
I_N	Index set of inequality constraints
I_L	Index set of linear inequality constraints
I	Index of all constraints defining \mathcal{R}
g_j	Gradient of j-th constraint
β_j	Gradient of j-th constraint at the point \hat{x}
\mathcal{R}	Quadratically constrained convex region
\mathcal{R}_k	The region defined by the constraints in the set $I \setminus \{k\}$
$\text{aff}(\mathcal{R})$	Smallest affine space containing \mathcal{R}
$\dim(\mathcal{R})$	Dimension of \mathcal{R}
$\text{int}(\ast)$	Interior of (\ast)
$\partial(\ast)$	Boundary of (\ast) ($\partial(\ast) = (\ast) \setminus \text{int}(\ast)$)
$I(x^\ast)$	The index set of constraints active at x^\ast
$N(A)$	Null space of A
$R(A)$	Range space of A
$C(A)$	Column space of A
$A(J)$	The matrix with columns a_j and B_j , $j \in J \subset I$
$\mathcal{B}(\bar{x}, \epsilon)$	A ball centered at \bar{x} , with radius ϵ
$\mathcal{R}(J)$	Region defined by the constraints with indices $J \subset I$
$\mathcal{R}(z)$	Truncated feasible region with truncation level z
$\mathcal{R}_z(J)$	Truncated feasible region defined by constraints $j \in J \subset I$

Symbol	Meaning
$\mathcal{M}(\bar{x})$	Linear manifold passing through \bar{x}
$F(x, z)$	Potential function at x with truncation level z
$\ x\ $	Euclidean norm of the vector x
\hat{x}^k	k-th element of the auxiliary sequence
$\nabla Q(x)$	Gradient of $Q(x)$
$\nabla^2 Q(x)$	Hessian matrix of $Q(x)$

CHAPTER 1

INTRODUCTION

1.1. Overview and Outline of Thesis.

This thesis is concerned with the problem of minimizing a convex quadratic function subject to convex quadratic constraints. This problem, which will be referred to as the QCQP problem, has linear programmes and quadratic programmes as special cases. The QCQP problem has received considerable attention in recent years [7.9.16.21.26.27.30.34.36.45]. Algorithms for the solution of the QCQP problem include interior point methods (Goldfarb et al. [21], Jarre [26], Mehrotra and Sun [34.36]), a cutting plane algorithm (Baron [2]), a reduced gradient method (Cole et al. [10]), a dual method (Ecker and Niemi [16]), a perturbation method (Fang and Rajasekera [18]), and a triangularization method (Phan-huy-Hao [44]).

The problem of quadratically constrained quadratic programming (QCQP) is significant because of its many applications which include location and production planning problems [44] and methods for the solution of nonlinear programming problems [55].

Although significant contributions have been made in the area of the QCQP problem, particularly in the application of the method of centers and of the barrier function method as solution algorithms for this problem, there are still many issues to be settled. The purpose of this thesis is to deal with some of these issues.

Our interest is in the minimal representation of the feasible region, in the problem of dealing with feasible regions that are unbounded or not full dimensional, and in methods for improving the rate of convergence.

Interior point methods have a long history. Early works include the monograph by Fiacco and McCormick [19], the paper on the method of centers by Huard [25]

and Dikin's interior point method for linear programmes [14]. Interest in interior point methods surged in 1984 when Karmarkar [28] published a linear programming algorithm, which proved to have very good practical performance as well as a good worst-case complexity bound. Since the publication of Karmarkar's paper many researchers, e.g., Den Hertog [11,12,13], Lustig [32], Gill et al. [20], Goldfarb [21], Gonzaga [22], Renegar [46], Sonnevend [49], Vaidya [53], have studied interior point methods for linear and quadratic programming. In [39] Nesterov and Nemirovskii developed a theory of interior point methods that applies to more general convex programming problems.

One of the concerns in solving a QCQP problem is that of obtaining a simplified representation of the quadratically constrained convex feasible regions. Finding such a simplified representation could lead to a significant reduction in the computational burden required to solve the problem, and could serve as a "preprocessor" for QCQP algorithms. In Chapter 2, we characterize minimal representations of quadratically constrained feasible regions. We define a representation to be minimal if every other representation has either more quadratic constraints, or has the same number of quadratic constraints and at least as many linear constraints. We will prove that a representation is minimal if and only if it contains no redundant constraints, no pseudo-quadratic constraints and no implicit equality constraints. We define a pseudo-quadratic constraint as a quadratic constraint that can be replaced by a finite number of linear constraints, and an implicit equality constraint as an inequality constraint which holds as an equality at all feasible points. In order to prove the minimal representation theorem, we present results on the faces of the quadratically constrained feasible region. (Faces for QCQP regions are analogous to facets for polyhedral sets.) The critical result is that if the boundaries of two quadratic constraints coincide on any open set, then they coincide everywhere. In Chapter 2 we also provide algorithms that can be used to detect quadratic implicit

equalities and pseudo-quadratic constraints. The redundant constraints can be identified by random methods, for example, by the hypersphere directions (HD) method [3].

In Chapter 3 we are concerned with the unboundedness problem. To date, methods for detecting unboundedness have, in fact, only dealt with the problem of detecting unbounded rays, that is, rays in the feasible region along which the objective function is unbounded. For example, this is the case for the methods in [7] which are concerned with unboundedness of both convex and concave QCQP problems. One of the intriguing differences between linear and quadratic programming problems and convex QCQP problems is that linear and quadratic programmes are unbounded if and only if there exists an unbounded ray, whereas QCQP problems can be unbounded even when no such ray exists. This was demonstrated by an example in Murty [37]. The same example appeared in the paper by Camerini et al. [4] and is given as Example 3.1 in this thesis. In Chapter 3 we show that although unboundedness of the QCQP problem does not imply the existence of an unbounded ray, the QCQP is unbounded if and only if there exists a related problem with fewer constraints and variables for which an unbounded ray exists. This result follows from the necessary and sufficient conditions for unboundedness of the QCQP. These conditions are given in the form of an algorithm, namely, Algorithm D, which, at each iteration, requires the identification of implicit equality constraints in a homogeneous linear system. Since the implicit equalities can be detected in a finite number of simplex steps, and since each iteration of the algorithm reduces the number of the constraints and the number of variables, the algorithm terminates in a finite number of steps.

In Chapter 4 we are concerned with the application of a method of centers to QCQP problems having a faulty feasible region. We say that a region is faulty if it is either unbounded or not full dimensional. Each iteration of a method of centers

must determine a center of a “truncated feasible region”. If the feasible region is faulty, then the truncated feasible region may have no interior, it may have no center, it may have an infinity of centers, or it may have a unique center. Thus, the difficulty caused by faulty feasible regions is that, even when the problem has a solution, a method of centers may fail.

Typically, faulty regions have been dealt with by the direct method of adding more constraints and variables, often involving a “Big M” constant. Our method is unique in that it results in a reduction in the number of constraints and in the number of variables. The method, which is called Algorithm E, is related to Algorithm D in Chapter 3.

In Chapter 5 we show how to accelerate the rate of convergence of a sequence $\{x^k\}$ generated by the application of a logarithmic barrier function method or a method of centers to the QCQP problem. We first show how the ideas in [29] can be generalized and used to construct an auxiliary sequence $\{\bar{x}^k\}$ that converges superlinearly faster than the original sequence $\{x^k\}$. We then show how the auxiliary sequence can be used to create a superlinearly convergent long step predictor-corrector method. Chapter 5 also presents the results of preliminary numerical testing. The results suggest that our method could compete successfully with other interior point methods.

Chapter 6 contains concluding remarks and future research directions. Most of the future research directions involve extensions of the results obtained in Chapters 2, 3, 4 and 5 to more general classes of convex or even nonconvex programmes. In particular, one such possibility involves an extension to semidefinite programming problems [45,54].

We end this chapter with a description of the problem; a discussion of both a method of centers and a logarithmic barrier function method, and a description of an algorithm for detecting implicit equalities in homogeneous linear systems.

1.2. Problem Definition and Background Material.

In this thesis we are concerned with the quadratically constrained convex quadratic program (QCQP)

$$\begin{aligned} \text{minimize: } & Q_0(x) = a_0^T x + \frac{1}{2} x^T B_0 x \\ \text{subject to: } & x \in \mathcal{R}, \end{aligned} \quad (1.2.1)$$

where the feasible region \mathcal{R} is given by

$$\mathcal{R} = \{x \in \mathbb{R}^n \mid Q_i(x) := a_i^T x + \frac{1}{2} x^T B_i x - b_i \leq 0, \ i \in I := \{1, \dots, m\}\}. \quad (1.2.2)$$

We assume that the matrices B_i are symmetric and positive semidefinite, and that $\mathcal{R} \neq \emptyset$. For the sake of convenience we will sometimes omit the subscript “0” for the quantities defining the objective function. We will also use the notation $I_0 = \{0, 1, \dots, m\}$, at which times we understand that $b_0 = z$ is some unspecified constant satisfying $z > Q_0(x^*)$, where x^* is a minimizer for QCQP, in case the problem is bounded.

We represent the smallest affine space containing \mathcal{R} by $\text{aff}(\mathcal{R})$. The dimension of \mathcal{R} is denoted by $\dim(\mathcal{R})$ and is defined to be the dimension of $\text{aff}(\mathcal{R})$. We use the symbols $\text{int}(\cdot)$ and $\partial(\cdot)$ to denote the interior and boundary, respectively, of the set (\cdot) . Note that $\partial(\mathcal{R}) = \mathcal{R} \setminus \text{int}(\mathcal{R})$.

Sufficient conditions that a point x^* be a solution to the QCQP problem defined in (1.2.1) - (1.2.2) is that there exists vector $u^* \in \mathbb{R}^m$, which, along with x^* , satisfies

$$Q_i(x^*) \leq 0, \quad i \in I, \quad (1.2.3.a)$$

$$u_i^* Q_i(x^*) = 0, \quad i \in I, \quad (1.2.3.b)$$

$$u_i^* \leq 0, \quad i \in I, \quad (1.2.3.c)$$

$$\nabla \mathcal{L}(x^*, u^*) = \nabla Q(x^*) - \sum_{i=1}^m u_i^* \nabla Q_i(x^*) = 0, \quad (1.2.3.d)$$

where

$$\mathcal{L}(x, u) = Q(x) - \sum_{i=1}^m u_i Q_i(x)$$

is the Lagrangian function for the QCQP. If, in addition to conditions (1.2.3), we assume that for every nonzero y satisfying

$$y^T \nabla Q_i(x^*) = 0, \quad \forall i \in D^* = \{ i \mid u_i^* < 0 \}$$

it follows that

$$y^T \nabla^2 \mathcal{L}(x^*, u^*) y > 0, \tag{1.2.4}$$

then x^* is a unique minimizer of the problem (1.2.1) - (1.2.2).

When the Kuhn-Tucker first-order constraint qualification ([19, page 19]), is satisfied, then conditions (1.2.3) are also necessary conditions for x^* to be a minimum of the QCQP problem.

1.2.1. A Logarithmic Barrier Function Method and a Method of Centers.

In this section we give a general description of two algorithms, a logarithmic barrier function method and a method of centers for the solution of the QCQP problem. Variations of these algorithms are in current use for the solution of the QCQP as well as other optimization problems [12,20,21,22,26,27]. They are also used throughout the thesis. While they are presented as separate algorithms, Lootsma [31] has shown that the method of centers is, in fact, a nonparametric barrier method. We present this relationship in Section 1.2.1.3. This relationship can be used to show that where the results in this thesis are given for one of the algorithms, they are also applicable to the other algorithm.

The choice of algorithm used to present a particular result depends upon which is more convenient for the presentation of the result. The common feature of these methods is that they both generate a sequence of points $\{x^k\} \in \text{int}(\mathcal{R})$, that con-

verges to an optimal solution. Thus, these methods are also called “interior point methods”.

1.2.1.1. A Logarithmic Barrier Function Method: Algorithm A.

This section gives a description of a logarithmic barrier function method as it is used in this thesis. First, we define the logarithmic barrier function as

$$L(x, r) = Q(x) - r \sum_{i=1}^m \ln(-Q_i(x)), \quad (1.2.5)$$

where $r > 0$ is a scalar parameter. The algorithm follows.

Algorithm A: A logarithmic barrier function method

```

 $r_0 > 0, \ k := 0$ 
repeat
    Set     $x^k = \operatorname{argmin}\{L(x, r_k)\}, \ r_{k+1} \in (0, r_k)$ 
    if     the convergence criteria are not satisfied
    then    $k := k + 1$ 
    else   stop := true
until    stop     $\{x^* \approx x^k\}$ 
□

```

The following theorem establishes convergence of Algorithm A.

Theorem 1.2.1 [19, Theorem 14]

Assume that the gradients $\nabla Q_i(x^*), i \in I(x^*) := \{i \in I \mid Q_i(x^*) = 0\}$ are linearly independent, that strict complementarity holds for $u_i^* Q_i(x^*) = 0, i \in I$, and that conditions (1.2.3.a) - (1.2.3d) are satisfied by (x^*, u^*) .

Then there is a neighbourhood about $r = 0$, for which there exists a unique, continuously differentiable function $[x(r), u(r)]$ satisfying

$$\nabla Q(x(r)) - \sum_{i \in I} u_i(r) \nabla Q_i(x(r)) = 0,$$

and

$$u_i(r) Q_i(x(r)) = r, \quad i \in I,$$

such that $x(r)$ ($r > 0$) describes a unique isolated trajectory of local minima of $L(x, r)$, $x(r) \rightarrow x^*$, and $u(r) \rightarrow u^*$.

The proof of Theorem 1.1.2 depends upon the facts that $L(x, r)$ is a strictly convex function on $\text{int}(\mathcal{R})$ - so that the $x^k, k > 0$, are unique - and that $L(x^k, r_k)$ is a monotonically decreasing sequence. (The latter property is a direct consequence of the definition of $L(x, r)$).

We note that if \mathcal{R} is unbounded, then $L(x, r)$ can be unbounded from below, in which case Algorithm A cannot be implemented. In Chapter 4 we show how to implement a modification of Algorithm A which can deal with unbounded feasible regions.

1.2.1.2. A Method of Centers: Algorithm B.

This section gives a description of a method of centers as it is used in this thesis. First, we define the truncated feasible region

$$\mathcal{R}(z) = \{ x \in \mathcal{R} \mid Q(x) < z \},$$

where z is a real scalar; and we define the potential function

$$F(x, z) = -m \ln(z - Q(x)) - \sum_{i=1}^m \ln(-Q_i(x)).$$

The algorithm follows.

Algorithm B: A Method of Centers

```

 $x^0 \in \text{int}(\mathcal{R}), \quad k = 0$ 
repeat
    for  $z_k = Q(x^k)$  find  $x^{k+1} = \text{argmin} \{ F(x, z_k) \mid x \in \text{int}(\mathcal{R}(z_k)) \}$ 
    if convergence criteria are not satisfied
        then  $k := k + 1$ 
    else stop := true
until stop  $\{x^* \approx x^k\}$ 
□

```

We refer to x^{k+1} as the center, or analytic center, of $\mathcal{R}(z_k)$.

Many variants of Algorithm B have been developed by researchers to solve QCQP. Among them are the short step and the long step path following methods. In the short step methods [26,34], a single Newton iteration is used to determine an

approximation to the new estimate x^{k+1} . That is, we set $x^{k+1} = x^k + \sigma_k p_k$, where the search direction p_k satisfies the Newton equation

$$\nabla^2 F(x^k, z_k) p_k = -\nabla F(x^k, z_k), \quad (1.2.6)$$

and where σ_k is some scaling factor (step size), which is a measure of the closeness to the optimal solution. Usually

$$\sigma_k = \frac{\beta}{\sqrt{p_k \nabla^2 F(x^k, z_k) p_k}},$$

where β is some small value parameter.

In the long step path following methods [12,46] several iterations of Newton's method, using the optimal step size, i.e.,

$$\sigma_k = \operatorname{argmin}_{\sigma > 0} F(x^k + \sigma p_k)$$

are used to find an estimate of x^{k+1} .

In either case, the truncation parameter z is updated by the formula

$$z_{k+1} = z_k + \theta(Q(x^k) - z_k),$$

for some parameter $\theta \in (0, 1)$ so that $x^{k+1} \in \mathcal{R}(z_k)$.

Explicit expressions for the gradient and Hessian of $F(x, z)$ are given by

$$\nabla F(x, z) = \frac{m \nabla Q_0(x)}{(z - Q_0(x))} + \sum_{j \in I} \frac{\nabla Q_j(x)}{(b_j - Q_j(x))}$$

and

$$\begin{aligned} \nabla^2 F(x, z) &= \frac{m \nabla Q_0(x) \nabla Q_0(x)^T}{(z - Q_0(x))^2} + \sum_{j \in I} \frac{\nabla Q_j(x) \nabla Q_j(x)^T}{(b_j - Q_j(x))^2} \\ &\quad + \frac{m B_0}{(z - Q_0(x))} + \sum_{j \in I} \frac{B_j}{(b_j - Q_j(x))}, \end{aligned}$$

respectively. We denote the column space and the null space of $\nabla^2 F(x, z)$ by $C(\nabla^2 F(x, z))$ and $N(\nabla^2 F(x, z))$, respectively. It is easy to see that

$$N(\nabla^2 F(x, z)) = N([a_0, a_1, \dots, a_m, B_0, B_1, \dots, B_m]). \quad (1.2.7)$$

Thus, both $N(\nabla^2 F(x, z))$ and $C(\nabla^2 F(x, z))$ are independent of x and z .

1.2.1.3. Remarks on Algorithms A and B.

We will first establish a relationship between Algorithm A and Algorithm B. Since $\nabla F(x, z_k) = 0$ at $x = x^{k+1}$, we have

$$\frac{m \nabla Q(x^{k+1})}{Q(x^{k+1}) - Q(x^k)} - \sum_{i=1}^m \frac{\nabla Q_i(x^{k+1})}{Q_i(x^{k+1})} = 0.$$

Since $\nabla L(x, r_{k+1}) = 0$ at $x = x^{k+1}$, we have

$$\nabla Q(x^{k+1}) - r_{k+1} \sum_{i=1}^m \frac{\nabla Q_i(x^{k+1})}{Q_i(x^{k+1})} = 0.$$

Defining

$$r_{k+1} = \frac{Q(x^{k+1}) - Q(x^k)}{m}$$

we get that

$$\operatorname{argmin}\{L(x, r_{k+1})\} = \operatorname{argmin}\{F(x, x^k) \mid x \in \operatorname{int}(\mathcal{R}(z_k))\},$$

or that the Algorithm A and Algorithm B would produce identical sequences $\{x^k\}$. This result was first given by Lootsma [31]. An interpretation of the result is that the truncation level $Q(x^k)$ in the method of centers acts like an implicit parameter “ r ” in the barrier method.

A difficulty with both Algorithm A and Algorithm B is that an initial interior point is needed to start the algorithms. (Note that the initial interior point is needed to solve for $x^0 = \operatorname{argmin}\{L(x, r_0)\}$.) Techniques used to determine such a point either require a repeated application of the method of itself [19], or require the addition of a new variable and a new constraint involving a big M constant [34,36]. In Chapter 4 we give a technique to provide an interior point given a feasible point that neither requires a Big M, nor an application of the method itself.

1.2.2. Implicit Equalities in Homogeneous Linear Systems.

In the remaining chapters, we will often require a method to determine implicit equalities in homogeneous linear systems. In particular, in Chapters 2, 3 and 4 the

method is used to detect and remove quadratic implicit equalities, to determine the boundedness of the QCQP, and to determine the existence of an analytic center of the QCQP, respectively. As previously mentioned, an implicit equality is an inequality that holds as equality for all solutions to the system. In this section we discuss the issue of how to determine implicit equalities.

For ease of presentation of the method we consider the linear system

$$a_i^T s \leq 0, \quad i \in J, \quad (1.2.8)$$

where J is a subset of I , and where the a_i are as defined in Section 1.2. Let's first determine whether or not $a_1^T s \leq 0$ is an implicit equality in (1.2.8). This problem is, of course, equivalent to the determination of whether or not System I or System II of Farkas' theorem of the alternative has a solution, where

$$\text{System I: } a_i^T x \leq 0, \quad i \in J \setminus \{1\}, \quad a_1^T s < 0,$$

$$\text{System II: } \sum_{i \in J \setminus \{1\}} \lambda_i a_i = a_1, \quad \lambda_i \leq 0, \quad i \in J \setminus \{1\}$$

It is easy to see that this is equivalent to the determination that either $s = 0$ is a solution to the LP

$$\min \{ a_1^T s \mid a_i^T s \leq 0, \quad i \in J \setminus \{1\} \},$$

or that the LP is unbounded from below. In the latter case, $a_1^T s \leq 0$ is not an implicit equality, while in the former case, it is an implicit equality. Solving this LP will require that we determine feasibility of the linear system

$$a_1 = \sum_{i \in J \setminus \{1\}} \lambda_i a_i, \quad \lambda_i \leq 0, \quad i \in J \setminus \{1\}.$$

Since we need to identify all implicit equalities in (1.2.8), we actually check for feasibility to all the systems

$$a_k = \sum_{i \in J \setminus \{k\}} \lambda_i a_i, \quad \lambda_i \leq 0, \quad i \in J \setminus \{k\}.$$

In some cases we want to identify implicit equalities in systems of the form

$$\begin{aligned} B_i s &= 0, \quad i \in J \\ a_i^T s &\leq 0, \quad i \in J. \end{aligned} \tag{1.2.9}$$

This system is dealt with as above, by first projecting the vectors $a_i, i \in J$ onto the intersection of $N(B_i), i \in J$.

There are two other issues to address. First, the same technique can be applied to linear systems of the form

$$g_i^T x \leq \beta_i, \quad i \in I(\bar{x})$$

where $g_i^T \bar{x} = \beta_i, i \in I(\bar{x})$. This is the case for Algorithm C in Chapter 2.

Finally, the technique is useful in the determination of whether or not there is a nonzero solution to systems of the form given in (1.2.9). This case occurs, for example, as a consequence of Theorem (3.1.1). This is the case if and only if $\bar{s} = 0$ is the only optimal solution to each of the following $2n$ linear programs

$$\min\{(s_i) \mid a_j^T s \leq 0, \quad B_j s = 0, \quad j \in J\}, \quad i = 1, \dots, n,$$

and

$$\min\{-(s)_i \mid a_j^T s \leq 0, \quad B_j s = 0, \quad j \in J\}, \quad i = 1, \dots, n,$$

where $(\bar{s})_i$ is the i -th component of vector \bar{s} . Note that, after using projections to eliminate the equalities, these LPs are of the form corresponding to System I and System II given above.

CHAPTER 2

MINIMAL REPRESENTATION

2.1. Introduction.

In this chapter we are concerned with the characterization of minimal representations of the feasible region \mathcal{R} . We first introduce our definition of the minimal representation and we then provide necessary and sufficient conditions for a representation to be minimal. In order to prove the minimal representation theorem, we prove the critical result that if the boundaries of two quadratic constraints coincide on any open set, then they coincide everywhere. Afterwards we present the steps of an algorithm to obtain the minimal representation. This algorithm depends upon our new methods, which will be presented in this chapter, to detect both implicit equalities and pseudo quadratic constraints.

As \mathcal{R} may be not full dimensional, a minimal representation of \mathcal{R} may well contain linear equality constraints that represent $\text{aff}(\mathcal{R})$. Thus, for convenience, we explicitly include linear equality constraints at the outset. Linear inequality constraints are included by setting the corresponding Hessian matrix to zero. We will use τ to denote the number of equality constraints, m to denote the number of linear inequality constraints and q to denote the number of quadratic inequality constraints. For the sake of presentation we redefine \mathcal{R} as follows:

$$\mathcal{R} = \{x \in \mathbb{R}^n \mid Ax = b; Q_i(x) \leq 0, i \in I_N\}, \quad (2.1.1)$$

where $A^T = [a_1, \dots, a_\tau]$, $b^T = [b_1, \dots, b_\tau]$, and $I_N = I_L \cup I_Q$, where $I_L = \{\tau + 1, \dots, \tau + \mu\}$ and $I_Q = \{\tau + \mu + 1, \dots, \tau + \mu + q\}$. The $Q_i(x)$ are as defined in Chapter 1.

We say that (2.1.1) is a representation of \mathcal{R} . Clearly, the region \mathcal{R} does not have a unique representation. For example, another representation can be given by

$$\mathcal{R} = \{x \in \mathbb{R}^n \mid \bar{A}x = \bar{b}; \bar{Q}_i(x) \leq 0, i \in I_{\bar{N}}\}, \quad (2.1.2)$$

where $\bar{A}^T = [\bar{a}_1, \dots, \bar{a}_\tau]$, $\bar{b}^T = [\bar{b}_1, \dots, \bar{b}_\tau]$, and $I_{\bar{N}} = I_{\bar{L}} \cup I_Q$, where $I_{\bar{L}} = \{\bar{\tau} + 1, \dots, \bar{\tau} + \bar{\mu}\}$, $I_Q = \{\bar{\tau} + \bar{\mu} + 1, \dots, \bar{\tau} + \bar{\mu} + \bar{q}\}$. The functions are given by $\bar{Q}_i(x) := \bar{a}_i^T x + 1/2 x^T \bar{B}_i x - \bar{b}_i$.

A natural question to ask is, “which representation is preferable, (2.1.1) or (2.1.2)?” Our preference is indicated by the following definition.

Definition 2.1.1. We say that (2.1.1) is a *minimal representation* of \mathcal{R} if every other representation, e.g., (2.1.2), has either more quadratic constraints ($\bar{q} > q$), or has the same number of quadratic constraints ($\bar{q} = q$) and at least as many linear constraints ($\bar{\tau} + \bar{\mu} \geq \tau + \mu$).

We will prove that (2.1.1) is a minimal representation of \mathcal{R} if and only if it contains no redundant constraints, no implicit equality constraints, and no pseudo-quadratic constraints. Consider the following definitions.

Definition 2.1.2. Constraint $k \in I_N$ is *redundant* with respect to the representation (2.1.1) if $\mathcal{R} = \mathcal{R}_k$, where

$$\mathcal{R}_k = \{x \in \mathbb{R}^n \mid Ax = b; Q_i(x) \leq 0, i \in I_N \setminus \{k\}\}.$$

Constraint $k \in \{1, \dots, \tau\}$ is *redundant* with respect to the representation (2.1.1) if $\mathcal{R} = \mathcal{R}_k$, where

$$\mathcal{R}_k = \{x \in \mathbb{R}^n \mid a_i^T x = b_i, i \in \{1, \dots, \tau\} \setminus \{k\} \mid Q_i(x) \leq 0, i \in I_N\}.$$

Definition 2.1.3. Constraint $k \in I_N$ is an *implicit equality* in \mathcal{R} if $Q_k(x) = 0$ whenever $x \in \mathcal{R}$.

Definition 2.1.4. Constraint $k \in I_Q$ is *pseudo-quadratic* with respect to the representation (2.1.1) if it is necessary (i.e., non-redundant) and if there exists a finite set of linear inequalities $P^k x \leq p^k$ such that

$$\mathcal{R} = \{x \in \mathbb{R}^n \mid Ax = b, Q_i(x) \leq 0, i \in I_N \setminus \{k\}, P^k x \leq p^k\}.$$

Example 2.1.1. Consider Figure 2.1(a). Constraint (1) is a linear equality, constraints (2) and (3) are linear inequalities and constraint (4) is a quadratic inequality. The feasible region represented by the constraints is the line segment inside the ellipsoid. The quadratic constraint is pseudo-quadratic since it can be replaced by the two linear inequalities (5) and (6), as shown in Figure 1(b). In Figure 2.1(b) we see that constraints (2) and (3) are redundant. The minimal representation of the feasible region is given in Figure 2.1(c) and consists of one linear equality and two linear inequalities.

□

The results presented in this chapter are a generalization of Telgen's minimal representation theory for linearly constrained feasible regions [50], which, in turn, is a generalization of an earlier paper by Eckhardt [17].

In Section 2.2 we present an algorithm for the identification of all implicit equalities. Furthermore, we show that both the linear and the quadratic implicit equalities can be replaced by explicit linear equality constraints. In Section 2.3 we show that pseudo-quadratic constraints are easily identified, and that each pseudo-quadratic constraint can be replaced by two linear inequalities. In Section 2.4 we present results on the faces of the feasible region \mathcal{R} . These results are needed for the proof of the minimal representation theorem which is given in Section 2.5. In Section 2.6 we give a step by step procedure for obtaining a minimal representation. Our procedure makes use of the Hypersphere Directions (HD) method [3] for the identification of the redundant constraints. The chapter concludes with Section 2.7.

We end this section with an introduction to some of the notation that will be used throughout this chapter.

The surface (boundary) of the inequality constraint $k \in I_N$ is given by

$$S_k := \{x \in \mathbb{R}^n | Q_k(x) = 0\}.$$

The face of \mathcal{R} associated with the k -th constraint is denoted by \mathcal{F}_k and is given by

$\mathcal{F}_k = S_k \cap \mathcal{R}$. The interior of \mathcal{F}_k , denoted by $\text{int}(\mathcal{F}_k)$, is given by

$$\text{int}(\mathcal{F}_k) = \{x \in \mathcal{F}_k \mid \exists \epsilon > 0 \ni \mathcal{B}(x, \epsilon) \cap S_k \subset \mathcal{F}_k\},$$

where

$$\mathcal{B}(x, \epsilon) := \{y \in \mathbb{R}^n \mid \|y - x\| \leq \epsilon\}.$$

As can be seen in Fig. 2.1(d), a face may have many components. We denote the i -th component of face \mathcal{F}_k by \mathcal{F}_k^i . A component is defined as follows. Let $\hat{x} \in \text{int}(\mathcal{F}_k)$. The component of \mathcal{F}_k containing \hat{x} is the set $\mathcal{F}_k^i \subseteq \mathcal{F}_k$ where $x \in \mathcal{F}_k^i$ if $x \in \mathcal{F}_k$ and if there exists a continuous arc $\{\phi(\alpha) \in \mathbb{R}^n \mid 0 \leq \alpha \leq 1\} \subset \text{int}(\mathcal{F}_k)$ with $\phi(0) = \hat{x}$ and $\phi(1) = x$.

For any $\hat{x} \in \mathcal{R}$, we denote the index set of all constraints active at \hat{x} by $I(\hat{x})$. That is, $Q_i(\hat{x}) = 0$ for $i \in I(\hat{x})$, and $Q_i(\hat{x}) < 0$ for $i \notin I(\hat{x})$.

2.2. Implicit Equality Constraints.

We begin with a useful characterization of an implicit equality constraint.

Lemma 2.2.1. Let $\hat{x} \in \mathcal{R}$ and $k \in I_N$. Constraint k is an implicit equality iff $Q_k(\hat{x}) = 0$; and $B_k s = 0$ and $a_k^T s = 0$ for all s with $\hat{x} + s \in \mathcal{R}$.

Proof. Let $g_k = a_k + B_k \hat{x}$ and consider any $(\hat{x} + s) \in \mathcal{R}$. If $Q_k(\hat{x}) = 0$, $B_k s = 0$ and $a_k^T s = 0$ then $Q_k(\hat{x} + s) = Q_k(\hat{x}) + g_k^T s + \frac{1}{2} s^T B_k s = 0$ so that constraint k is an implicit equality. Now suppose that constraint k is an implicit equality. We have $Q_k(\hat{x}) = 0$ and $Q_k(\hat{x} + s) = 0$. Since \mathcal{R} is convex, we also have that $\hat{x} + \sigma s \in \mathcal{R}$ for $0 \leq \sigma \leq 1$. Thus, for all $0 \leq \sigma \leq 1$, we have

$$Q_k(\hat{x} + s) = Q_k(\hat{x}) = Q_k(\hat{x} + \sigma s) = 0.$$

This, together with the fact that B_k is positive semidefinite yields that $B_k s = 0$, which, together with $Q_k(\hat{x} + s) = 0$, gives $a_k^T s = 0$. \square

The following corollary to Lemma 2.2.1 shows that if constraint k is an implicit equality constraint it can be replaced by the linear system $A^k x = b^k$, where

$$A^k = \begin{bmatrix} B_k \\ a_k^T \end{bmatrix}, \quad b^k = \begin{bmatrix} B_k \hat{x} \\ a_k^T \hat{x} \end{bmatrix},$$

and where \hat{x} is any point in \mathcal{R} .

Corollary 2.2.1. Let $\hat{x} \in \mathcal{R}$ and suppose that constraint k is an implicit equality. Then

$$\mathcal{R} = \mathcal{R}^k := \{x \in \mathbb{R}^n \mid Ax = b, Q_i(x) \leq 0, i \in I_N \setminus \{k\}, A^k x = b^k\}.$$

Proof. We will show that $\mathcal{R} \subset \mathcal{R}^k$ and $\mathcal{R}^k \subset \mathcal{R}$. Clearly, we need only consider constraint $Q_k(x) \leq 0$ defining \mathcal{R} and constraint $A^k x = b^k$ defining \mathcal{R}^k . Let $x \in \mathcal{R}$ and let s be such that $x = \hat{x} + s$. Since constraint k is an implicit equality it follows from Lemma 2.2.1 that $B_k s = 0$ and $a_k^T s = 0$. Thus, $A^k(\hat{x} + s) = A^k \hat{x} = b^k$, and $x \in \mathcal{R}^k$ so that $\mathcal{R} \subset \mathcal{R}^k$.

Now let $x \in \mathcal{R}^k$. We can write $x = \hat{x} + s$ and note that $A^k(\hat{x} + s) = b^k$ implies that $A^k s = 0$, i.e., that $B_k s = 0$ and $a_k^T s = 0$. We have

$$\begin{aligned} Q_k(x) &= Q_k(\hat{x} + s) \\ &= Q_k(\hat{x}) + (a_k + B_k \hat{x})^T s + \frac{1}{2} s^T B_k s \\ &= 0. \end{aligned}$$

Thus, $x \in \mathcal{R}$ and $\mathcal{R}^k \subset \mathcal{R}$. \square

Example 2.2.1 Consider the system given by $x_3 = 0$ and $x_2^2 - x_3 \leq 0$. The feasible region, which can be seen in Figure 2.2, is the spine of the trough given by the quadratic constraint. The quadratic is an implicit equality and, according to Corollary 2.2.1, it can be replaced by the linear constraints $x_2 = 0$ and $x_3 = 0$. Note that the second constraint is redundant.

\square

Clearly, we need at most $(n - \dim(\mathcal{R}))$ equalities to define $\text{aff}(\mathcal{R})$. Suppose that there are p implicit quadratic equalities. It then follows from Corollary 2.2.1 that we would add $(n+1)p$ equality constraints to the existing r equalities. Thus, $[(n+1)p + r - n + \dim(\mathcal{R})]$ of the equalities would be redundant. We note that redundant linear equalities are relatively easy to remove using, for example, Gaussian elimination.

The next requirement is a method to detect the complete set of implicit equality constraints. We propose the following Algorithm.

Algorithm C: A Method to Remove Quadratic Implicit Equalities

```

 $\hat{x} \in \mathcal{R}, \hat{J}_0 = I(\hat{x}), k = 0$ 
repeat
  for  $j \in \hat{J}_k$  set  $g_j = a_j + B_j \hat{x}, \beta_j = g_j^T \hat{x}$ 
  Find the set  $\hat{I}_k$  of all implicit equalities in
   $\{g_j^T x \leq \beta_j, j \in \hat{J}_k; B_j x = B_j \hat{x}, a_j^T x = a_j^T \hat{x}, j \in \hat{J}_0 \setminus \hat{J}_k\}$ 
  if  $\hat{I}_k \neq \emptyset$  then
     $\hat{J}_{k+1} := \hat{J}_k \setminus \hat{I}_k$ 
     $k := k + 1$ 
  else stop := true
until stop { Constraints  $i, i \in \hat{J}_0 \setminus \hat{J}_k$  are implicit equalities }
□

```

In order to determine the set \hat{I}_k , each step of Algorithm C requires the use of the method discussed in Section 1.2.2. The justification of the Algorithm is found in the following two lemmas.

Lemma 2.2.2. Let $\hat{x} \in \mathcal{R}$ and for each $i \in I(\hat{x})$ define $g_i = a_i + B_i \hat{x}$ and $\beta_i = g_i^T \hat{x}$. If the constraint $g_k^T x \leq \beta_k$ is an implicit equality in the system $\{g_i^T x \leq \beta_i, i \in I(\hat{x})\}$ then the constraint $Q_k(x) \leq 0$ is an implicit equality in the system $\{Q_i(x) \leq 0, i \in I(\hat{x})\}$.

Proof. Let $\mathcal{R}_L = \{x \in \mathbb{R}^n \mid g_i^T x \leq \beta_i, i \in I(\hat{x})\}$ and let $\mathcal{R}_Q = \{x \in \mathbb{R}^n \mid Q_i(x) \leq 0, i \in I(\hat{x})\}$. Note that, since \mathcal{R}_Q is convex, it follows that $\mathcal{R}_Q \subset \mathcal{R}_L$. Since the constraint $g_k^T x \leq \beta_k$ is an implicit equality in the system $g_i^T x \leq \beta_i, i \in I(\hat{x})$ it follows that $g_k^T x = \beta_k$ for all $x \in \mathcal{R}_L$. We need only show that $Q_k(x) = 0$ for all $x \in \mathcal{R}_Q$. Since $g_k^T x = \beta_k$ for all $x \in \mathcal{R}_L$ and since $\mathcal{R}_Q \subset \mathcal{R}_L$ it follows that $g_k^T x = \beta_k$ for all $x \in \mathcal{R}_Q$. Let $\bar{x} \in \mathcal{R}_Q$ be arbitrary but fixed. We have that

$$\begin{aligned}
Q_k(\bar{x}) &= Q_k(\hat{x} + (\bar{x} - \hat{x})) \\
&= Q_k(\hat{x}) + g_k^T (\bar{x} - \hat{x}) + \frac{1}{2} (\bar{x} - \hat{x})^T B_k (\bar{x} - \hat{x}) \\
&= \frac{1}{2} (\bar{x} - \hat{x})^T B_k (\bar{x} - \hat{x}) \\
&\geq 0 \quad .
\end{aligned}$$

Since $\bar{x} \in \mathcal{R}_Q$ we also have $Q_k(\bar{x}) \leq 0$. Thus, $Q_k(x) = 0$ for all $x \in \mathcal{R}_Q$. \square

Lemma 2.2.3. Let $\hat{x} \in \mathcal{R}$ and let \hat{I}_0 be as defined in Algorithm C. If $\hat{I}_0 = \emptyset$ there are no implicit equality constraints in \mathcal{R} .

Proof. First, suppose that $\hat{x} \in \mathcal{R}$ is such that $I(\hat{x}) = \emptyset$. Clearly, this implies that there are no implicit equalities. Now suppose that $I(\hat{x}) \neq \emptyset$. Since $\hat{I}_0 = \emptyset$ it follows that there exists a vector s such that $g_i^T(\hat{x} + s) < \beta_i$, $i \in I(\hat{x})$. This implies that $g_i^T s < 0$, $i \in I(\hat{x})$. For $\sigma > 0$ we have

$$\begin{aligned} Q_i(\hat{x} + \sigma s) &= Q_i(\hat{x}) + \sigma g_i^T s + \frac{\sigma^2}{2} s^T B_i s \\ &= \sigma g_i^T s + \frac{\sigma^2}{2} s^T B_i s. \end{aligned}$$

Since $g_i^T s < 0$ then for σ sufficiently small we have $Q_i(\hat{x} + \sigma s) < 0$, $i \in I(\hat{x})$. For $i \in I \setminus I(\hat{x})$ we have $Q_i(\hat{x} + \sigma s) < 0$ for σ sufficiently small since $Q_i(\hat{x}) < 0$, $i \in I \setminus I(\hat{x})$. This implies that there are no implicit equalities. \square

We note that if there are no quadratic constraints, i.e., if $q = 0$, then Algorithm C will terminate after one iteration. However, if $q > 0$, then the Algorithm C may require several iterations. This happens whenever the region $\{x \in \mathbb{R}^n \mid g_j^T x \leq 0, j \in \hat{J}_k\}$ has the dimension greater than $\dim(\mathcal{R})$. Consider the following example.

Example 2.2.1. The feasible region \mathcal{R} is defined by the system $Q_1(x) = x_2^2 - x_3 \leq 0$, $Q_2(x) = x_2^2 + x_3 \leq 0$ and $Q_3(x) = -x_2 \leq 0$. The boundaries of the constraints are shown in Figure 2.4. The first iteration of Algorithm C determines that the constraints $g_1^T x \leq \beta_1$ and $g_2^T x \leq \beta_2$ are implicit equalities with respect to $g_i x \leq \beta_i$, $i = 1, 2, 3$. Although inequality $g_3^T x \leq \beta_3$ is not an implicit equality, the corresponding quadratic constraint, $Q_3(x) \leq 0$, is an implicit equality with respect to $Q_i(x) \leq 0$, $i = 1, 2, 3$, and it is detected as such in the second iteration of Algorithm C. \square

We end this section with the following three lemmas. These lemmas, which deal with implicit equalities, will be required to complete the proof of the minimal representation theorem in Section 2.5.

Lemma 2.2.4. If the linear constraint $a_k^T x \leq b_k$, $k \in I_L$ is an implicit equality and if there are no equality constraints, i.e., if $\tau = 0$, then there is at least one more implicit equality.

Proof. Suppose that constraint $k \in I_L$ is the only implicit equality. Since $\tau = 0$, that there exists a point $\bar{x} \in \mathcal{R}$ such that $Q_k(\bar{x}) = 0$ and $Q_i(\bar{x}) < 0$ for all $i \in I_N \setminus \{k\}$. Consider the set of points $x(t) = \bar{x} - ta_k$. Since \bar{x} is an interior point of \mathcal{R}_k there exists an $\epsilon > 0$ such that $x(t)$ is in the interior of \mathcal{R}_k for all $0 \leq t \leq \epsilon$. Further,

$$Q_k(x(t)) = a_k^T \bar{x} - t \|a_k\|^2 - b_k < 0, \forall t > 0.$$

Thus, the point $x(\epsilon)$ is in the interior of \mathcal{R} . This contradicts k being an implicit equality. Thus, there must be at least one more implicit equality. \square

Lemma 2.2.5. If constraint $k \in I_L$ is the only implicit equality in (1.1), then constraint k is redundant.

Proof. Follows directly from Lemma 2.2.4 in Telgen [50]. \square

Lemma 2.2.6. Let \hat{I} be the set of indices of all implicit equalities in (1.1). If $\hat{I} \subset I_L$, and if for each $i \in \hat{I}$ we replace " $a_i^T x \leq b_i$ " with " $a_i^T x = b_i$ ", then the resulting system contains at least one redundant constraint.

Proof. Suppose that the number of linear constraints $\tau \geq 1$. Choose $k \in \hat{I}$ and replace all inequalities with indices in $\hat{I} \setminus \{k\}$ by equalities. Lemma 2.2.5 then implies that constraint k is redundant. Now suppose that $\tau = 0$. Lemma 2.2.4 ensures that there are at least two implicit equalities. Replace all but one of them with equalities to get $\tau > 0$. Lemma 2.2.5 implies that the remaining implicit equality is redundant. \square

2.3. Pseudo-Quadratic Constraints.

In this section we investigate properties of pseudo-quadratic constraints. For ease of presentation, we assume that there are no implicit or explicit equality constraints. After all, the implicit equalities can be detected and replaced by explicit equalities;

and explicit equalities are easily accommodated by considering the statements to be restricted to the appropriate linear manifold. We will show that pseudo-quadratic constraints are easily detected, and that each pseudo-quadratic constraint can be replaced by two linear inequality constraints.

Theorem 2.3.1. Constraint $k \in I_Q$ is pseudo-quadratic iff $a_k \in R(B_k)$ and $\text{rank}(B_k) = 1$.

Proof. Assume that constraint k is pseudo quadratic. We will first show that $\text{rank}(B_k) = 1$. Assume otherwise, i.e., that B_k has more than one nonzero eigenvalue. Let V be the eigenspace corresponding to the nonzero eigenvalues. Since the constraint $Q_k(x) \leq 0$ is positive definite on V and since there are no implicit or explicit equalities, the constraint $Q_k(x) \leq 0$ cannot be replaced by a finite number of linear constraints. This contradicts $Q_k(x) \leq 0$ being pseudo-quadratic. Thus, $\text{rank}(B_k) = 1$ and we can write $B_k = uu^T$ for some vector u . We can also write $a_k = \alpha u + z$ for some $z \in N(B_k)$ and $\alpha \in \mathbb{R}$. Since $z \in N(B_k)$ we have $z^T u = 0$. Since k is pseudo-quadratic it can be replaced by a finite number of linear inequalities. Since there are no implicit or explicit equalities each of those constraints corresponds to an $(n-1)$ dimensional face of \mathcal{R} . Let $a^T x \leq b$ be such an inequality and let x_0 be a point in the interior of the corresponding face.

Suppose that $z \neq 0$. Let v_1, \dots, v_{n-1} be a basis for subspace orthogonal to the vector a . For each i there exist a scalar $t_i > 0$ such that

$$\begin{aligned} Q_k(x_0 + tv_i) &= 0, \quad 0 \leq t \leq t_i \\ \iff (x_0^T uu^T v_i + \alpha u^T v_i + z^T v_i) + \frac{t}{2}(v_i^T u)^2 &= 0, \quad 0 \leq t \leq t_i, \\ \Rightarrow v_i^T u &= 0. \end{aligned} \tag{2.3.1}$$

Substituting $v_i^T u = 0$ in (2.3.1) gives $z^T v_i = 0, \forall i$. Since $u^T v_i = z^T v_i = 0$, for $i = 1, \dots, n-1$, it follows that $z = \gamma u$, for some scalar $\gamma \neq 0$. This contradicts $z^T u = 0$ and implies that $z = 0$. Thus, $a_k = \alpha u$. Now assume that $a_k \in R(B_k)$ and $\text{rank}(B_k) = 1$.

Since $\text{rank}(B_k) = 1$ and $a_k \in R(B_k)$, we can write $B_k = uu^T$ and $a = \alpha u$. Thus,

$$\begin{aligned} Q_k(x) \leq 0 &\iff \frac{1}{2}(x^T u)^2 + \alpha(x^T u) - b_k \leq 0 \\ &\iff u^T x \leq r_1, \quad u^T x \geq r_2. \end{aligned}$$

where $r_1 = -\alpha + \sqrt{\alpha^2 + 2b_k}$ and $r_2 = -\alpha - \sqrt{\alpha^2 + 2b_k}$. Note that $r_1 \neq r_2$ since $\alpha^2 + 2b_k = 0$ implies that constraint k is an implicit equality constraint, which contradicts our assumption. \square

Example 2.3.1. We consider the constraint $x^2 - 4x + 3 \leq 0$ with $b = -3$, $B = [2]$, $u = \sqrt{2}$, and $\alpha = -2\sqrt{2}$. Thus, $r_1 = 3\sqrt{2}$ and $r_2 = \sqrt{2}$, and the constraint can be replaced with $x \leq 3$ and $x \geq 1$.

\square

Example 2.3.2. We consider three examples. For Example 3.2(a) we consider the constraint $(x_1 - 2)^2 + (x_2 - 2)^2 - 1 \leq 0$ which defines a feasible region that is a circle centered at $(2, 2)$ with radius one. Clearly, this circular region cannot be described by a finite number of linear inequalities, and is not pseudo quadratic. Note that the Hessian is rank two. For Example 2.3.2(b) we consider the constraint $x_2^2 + x_1 - 2 \leq 0$ which defines a feasible region that is a parabola with vertex at $(2, 0)$ intersecting the x_2 axis at the points $(0, \pm\sqrt{2})$. Clearly, the parabolic region cannot be defined by a finite number of linear inequalities, and is not pseudo quadratic. While the Hessian matrix is rank one, its range space does not contain the vector $a = (0, 1)^T$. For Example 3.2(c) we consider the constraint $x_2^2 - 1 \leq 0$, which defines a feasible region that is a horizontal strip symmetric about the x_1 axis. This constraint is pseudo-quadratic and can be replaced by the constraints $x_2 \leq 1$ and $-x_2 \leq 1$. Note that the Hessian is rank one and that its range space contains the vector $a = (0, 0)^T$.

\square

Example 2.3.3 In this example, we have an explicit equality constraint. Consider the system given by $x_3 = 1$ and $x_2^2 - x_3 \leq 0$. In Figure 3, we see that the feasible region in \mathbb{R}^3 is the strip of the plane $x_3 = 1$ inside the trough defined by the

quadratic constraint. The quadratic constraint is a pseudo-quadratic constraint and it can be replaced by the linear constraints $-1 \leq x_2 \leq 1$.

We end this section with the observation that the objective function of a QCQP could be pseudo-quadratic. If the QCQP solution is not in the interior of \mathcal{R} , which can be checked by solving an unconstrained quadratic programme, then the QCQP can be solved by solving two QCQP problems with linear objectives. The two problems are minimize $\{u^T x | x \in \mathcal{R}\}$ and maximize $\{u^T x | x \in \mathcal{R}\}$. The optimal QCQP solution corresponds to the LP with the smaller objective function value. This observation may be of particular interest when \mathcal{R} is given by linear constraints.

□

2.4. Faces of \mathcal{R} .

In this section we prove results on the faces of \mathcal{R} . These results will be required for the proof of the main theorem. The key results in this section are contained in Theorem 2.4.1 and Theorem 2.4.2. As in Section 2.3, we assume that there are no implicit or explicit equality constraints.

Definition 2.4.1. The face \mathcal{F}_k is a *full dimensional face* if $\text{int}(\mathcal{F}_k) \neq \emptyset$.

From this definition it follows that \mathcal{F}_k is full dimensional iff it has positive $(n-1)$ dimensional Lebesgue measure. We define $\dim(\mathcal{F}_k)$ as the dimension of the smallest linear manifold containing (\mathcal{F}_k) . If there are no pseudo-quadratic constraints, then each full dimensional face corresponds to either a linear constraint with $\dim(\mathcal{F}_k) = n-1$, or to a quadratic constraint, with $\dim(\mathcal{F}_k) = n$. The following Lemma shows that the converse is also true.

Lemma 2.4.1. Suppose that \mathcal{F}_k is full dimensional and that constraint k is not pseudo-quadratic. If $\dim(\mathcal{F}_k) = n-1$, then $k \in I_L$. If $\dim(\mathcal{F}_k) = n$, then $k \in I_Q$.

Proof. Suppose that $\dim(\mathcal{F}_k) = n-1$ so that there exists n vectors y_1, \dots, y_n in

\mathcal{F}_k such that

$$L = \{x \mid x = \sum_{i=1}^n \alpha_i y_i, \sum_{i=1}^n \alpha_i = 1\}$$

has dimension $(n - 1)$. Since \mathcal{F}_k is full dimensional there exists an $x \in \text{int}(\mathcal{F}_k)$. If $x \notin L$ then $\dim(\mathcal{F}_k) = n$, which is a contradiction. Thus, $x \in L$ and $\mathcal{F}_k \subset L$, and since constraint k is not a pseudo-quadratic constraint this implies that $k \in I_L$. Now suppose that $\dim(\mathcal{F}_k) = n$. Since $k \notin I_L$, then $k \in I_Q$. \square

Lemma 2.4.2. Suppose that \mathcal{F}_k is full dimensional and that constraint k is not pseudo-quadratic. We have that $\dim(\mathcal{F}_k) = n$ iff for all $x \in \text{int}(\mathcal{F}_k)$ and for all $\epsilon > 0$ there exist vectors y_1, \dots, y_n in $\mathcal{F}_k \cap \mathcal{B}(x, \epsilon)$ together with non negative scalars $\alpha_1, \dots, \alpha_n$ satisfying $\sum_{i=1}^n \alpha_i = 1$ such that $Q_k(\bar{x}) < 0$, where $\bar{x} = \sum_{i=1}^n \alpha_i x_i$.

Proof. We prove the forward implication using contradiction. Suppose that $\dim(\mathcal{F}_k) = n$. Since \mathcal{F}_k is full dimensional it follows from Lemma 4.1 that $k \in I_Q$. Suppose that there exists an $\hat{x} \in \text{int}(\mathcal{F}_k)$ and an $\hat{\epsilon} > 0$ such that, for all $y_1, \dots, y_n \in \mathcal{F}_k \cap \mathcal{B}(\hat{x}, \hat{\epsilon})$ and for all $\alpha_i \geq 0, i = 1, \dots, n$ with $\sum_{i=1}^n \alpha_i = 1$, we have $Q_k(\bar{x}) = 0$, where $\bar{x} = \sum_{i=1}^n \alpha_i y_i$. This implies that $\dim(\mathcal{F}_k \cap \mathcal{B}(\hat{x}, \hat{\epsilon})) = (n - 1)$. We note that $\mathcal{F}_k \cap \mathcal{B}(\hat{x}, \hat{\epsilon}) \subset L$, where L is the full dimensional linear manifold

$$L = \{x \mid x = \sum_{i=1}^n \alpha_i y_i, \sum_{i=1}^n \alpha_i = 1\}.$$

We will show that $L \subset S_k$. Let $x \in L$. We can write $x = y_1 + \sum_{i=2}^n \beta_i (y_i - y_1)$. For all $i = 1, \dots, n$ we have $Q_k(y_i) = 0$ and for all $i = 2, \dots, n$ we have $Q_k((y_i + y_1)/2) = 0$. Thus, for all $i = 2, \dots, n$ we have

$$Q_k(y_i) = Q_k(y_1) = Q_k((y_i + y_1)/2) = 0,$$

from which, together with the fact that each B_i is positive semidefinite, it follows that for $i = 2, \dots, n$ we have $B_k(y_i - y_1) = 0$ and $a_k^T(y_i - y_1) = 0$. Thus, $B_k s = 0$ and $a_k^T s = 0$, where $s = \sum_{i=2}^n \beta_i (y_i - y_1)$. We now have $Q_k(x) = Q_k(y_1) + (a_k + B_k y_1)^T s + (\frac{1}{2}) s^T B_k s = 0$ so that $x \in S_k$. Since constraint k is not pseudo-quadratic

and since $\dim(\mathcal{F}_k) = n$ there must exist a full dimensional component \mathcal{F}_k^j of \mathcal{F}_k with $\mathcal{F}_k^j \cap \mathcal{B}(\hat{x}, \hat{\epsilon}) = \emptyset$, which cannot be replaced by a linear constraint. This implies that there exists a point $\tilde{x} \in \text{int}(\mathcal{F}_k^j)$ such that the supporting hyperplane of $\{x \in \mathbb{R}^n \mid Q_k(x) \leq 0\}$ at \tilde{x} intersects L . This contradicts the convexity of the region $\{x \in \mathbb{R}^n \mid Q_k(x) \leq 0\}$ and the fact that $L \subset S_k$.

We now prove the backwards implication. The hypothesis implies that $\dim(\mathcal{F}_k) \neq (n - 1)$. Since \mathcal{F}_k is full dimensional, then $\dim(\mathcal{F}_k) = n$. \square

The following corollary shows that for full dimensional faces, all full dimensional components have the same dimension.

Corollary 2.4.1. If \mathcal{F}_k^i and \mathcal{F}_k^j are two distinct full dimensional components of the full dimensional face \mathcal{F}_k , then $\dim(\mathcal{F}_k^i) = \dim(\mathcal{F}_k^j)$. Furthermore, if constraint k is not pseudo-quadratic, then $\dim(\mathcal{F}_k^i) = \dim(\mathcal{F}_k^j) = n$.

Proof. Since constraint k can only be quadratic (linear faces have only one component) then the comments preceding Lemma 2.4.1 establish that if constraint k is not pseudo-quadratic then the components have dimension n . If constraint k is pseudo-quadratic all components have dimension $(n - 1)$. The second part of the proof follows from Lemma 2.4.2, since, in the proof of Lemma 2.4.2, we deal with $\mathcal{F}_k \cap \mathcal{B}(\hat{x}, \hat{\epsilon})$ rather than with \mathcal{F}_k . \square

For the remainder of this paper when we refer to face \mathcal{F}_k , we mean the union of all components of \mathcal{F}_k . We now show that under the assumptions that there are no implicit equality constraints, all faces of necessary constraints are full dimensional.

Lemma 2.4.3. Suppose that (2.1.1) contains no implicit equality constraints. If constraint $i \in I_N$ is necessary then face \mathcal{F}_i is full dimensional.

Proof. Since there are no implicit equalities, we need only show that $\text{int}(\mathcal{F}_i) \neq \emptyset$. Since all constraints are convex and since there are no implicit equalities, it can be shown that if constraint i is necessary then $\mathcal{R}_i \setminus \mathcal{R}$ is full dimensional. Thus, there exists a point $x_i \in \mathcal{R}_i$ with $Q_i(x_i) > 0$ and an $\epsilon_i > 0$ such that $\mathcal{B}(x_i, \epsilon_i) \subset \mathcal{R}_i \setminus \mathcal{R}$.

Since there are no implicit equality constraints, there exists a point $z_2 \in \text{int}(\mathcal{R})$. Let C_i be the convex hull of $\{z_2\} \cup B(x_i, \epsilon_i)$. Since $B(x_i, \epsilon_i) \subset \mathcal{R}_i \setminus \mathcal{R}$ and since $z_2 \in \text{int}(\mathcal{R})$ it follows that $C_i \cap S_i \subset \mathcal{F}_i$, and that $\text{int}(C_i) \cap S_i \neq \emptyset$. \square

We now want to prove that if two surfaces S_j and S_k match on any open set, then they match everywhere. Before we prove that result, which is contained in Theorem 2.4.1, we must prove the next lemma.

Lemma 2.4.4. Assume that constraint k is not pseudo-quadratic and that there are no implicit equalities. Let \mathcal{H} represent an $(n-1)$ dimensional hyperplane in \mathbb{R}^n . Let $\bar{x} \in \mathcal{F}_k$ and let $\epsilon > 0$ be arbitrary but fixed. If $\mathcal{F}_k \cap B(\bar{x}, \epsilon)$ is full dimensional with dimension n , then $\mathcal{F}_k \cap B(\bar{x}, \epsilon) \setminus \mathcal{H}$ is also full dimensional with dimension n .

Proof. Since $\dim(\mathcal{F}_k \cap B(\bar{x}, \epsilon)) = n$ and $\dim(\mathcal{H}) = n-1$, there exists an $\hat{x} \in \text{int}(\mathcal{F}_k \cap B(\bar{x}, \epsilon))$ and an $\hat{\epsilon} > 0$ such that $\mathcal{H} \cap B(\hat{x}, \hat{\epsilon}) = \emptyset$. Thus, $\hat{x} \in \text{int}(\mathcal{F}_k \cap B(\bar{x}, \epsilon) \setminus \mathcal{H})$ and $\mathcal{F}_k \cap B(\bar{x}, \epsilon) \setminus \mathcal{H}$ is full dimensional.

Since $\mathcal{F}_k \cap B(\bar{x}, \epsilon)$ is full dimensional with dimension n , then \mathcal{F}_k is full dimensional with dimension n . Since constraint k is not pseudo-quadratic it follows from Lemma 2.4.2 that there exist points $x_i \in \mathcal{F}_k \cap B(\hat{x}, \hat{\epsilon})$, $i = 1, \dots, n$, together with scalars $\alpha_i \geq 0$, $i = 1, \dots, n$, with $\sum_{i=1}^n \alpha_i = 1$ such that $Q_k(\bar{x}) < 0$, where $\bar{x} = \sum_{i=1}^n \alpha_i x_i$. Thus, $\dim(\mathcal{F}_k \cap B(\bar{x}, \epsilon) \setminus \mathcal{H}) = n$. \square

Theorem 2.4.1. Suppose that (2.1.1) contains no implicit equalities and no pseudo-quadratic constraints. Let $\bar{x} \in \text{int}(\mathcal{F}_k)$ and let $\epsilon > 0$ be such that $S_k \cap B(\bar{x}, \epsilon) \subset \mathcal{F}_k$. If $B(\bar{x}, \epsilon) \cap S_k = B(\bar{x}, \epsilon) \cap S_j$, $j \neq k$, then $S_k = S_j$.

Proof. Since $\text{int}(\mathcal{F}_k) \neq \emptyset$ it follows that \mathcal{F}_k is full dimensional. First suppose that $j, k \in I_L$. Since \mathcal{F}_k is full dimensional with $\dim(\mathcal{F}_k) = n-1$ and since $B(\bar{x}, \epsilon) \cap S_j = B(\bar{x}, \epsilon) \cap S_k$, where $\bar{x} \in \text{int}(\mathcal{F}_k)$, it easily follows that $S_k = S_j$.

Now suppose that $k \in I_Q$ and $j \in I_L$. Since \mathcal{F}_k is full dimensional and since there are no pseudo-quadratic constraints then $\dim(\mathcal{F}_k) = n$. It now follows from Lemma 2.4.2 that $\dim(B(\bar{x}, \epsilon) \cap S_k) = n$, which in turn implies that $\dim(B(\bar{x}, \epsilon) \cap S_j) = n$,

contradicting $j \in I_L$. So we cannot have $k \in I_Q$ and $j \in I_L$, and, likewise, we cannot have $j \in I_Q$ and $k \in I_L$.

We now consider the case when $k \in I_Q$ and $j \in I_Q$. Suppose that the constraints have been scaled so that

$$e_1^T B_j e_1 = e_1^T B_k e_1 = 1. \quad (2.4.1)$$

Further, since

$$[B_j | a_j] \neq [B_k | a_k]. \quad (2.4.2)$$

else the theorem is trivial, we can also assume, without loss of generality that $e_1^T [B_j | a_j] \neq e_1^T [B_k | a_k]$.

For all $x \in S_k \cap \mathcal{B}(\bar{x}, \epsilon)$ we have $Q_j(x) = Q_k(x) = 0$. Equating $Q_j(x)$ and $Q_k(x)$, along with equations (2.4.1) and (2.4.2) imply that

$$x_1 = F(x_2, \dots, x_n) = \frac{Q_j(0, x_2, \dots, x_n) - Q_k(0, x_2, \dots, x_n)}{e_1^T (B_k - B_j)(0, x_2, \dots, x_n)^T + (a_k)_1 - (a_j)_1}$$

provided that $x \notin \mathcal{H}$, where \mathcal{H} is the hyperplane given by the equation

$$e_1^T (B_k - B_j)(0, x_2, \dots, x_n)^T + (a_k)_1 - (a_j)_1 = 0$$

(Note that x_1 is not equal to zero for arbitrary x_2, \dots, x_n .)

Since \mathcal{F}_k is full dimensional and since $\dim(S_k \cap \mathcal{B}(\bar{x}, \epsilon)) = n$, it follows from Lemma 2.4.4 that $S_k \cap \mathcal{B}(\bar{x}, \epsilon) \setminus \mathcal{H}$ is full dimensional with dimension n . Therefore, there exists an $\hat{x} \in S_k \cap \mathcal{B}(\bar{x}, \epsilon) \setminus \mathcal{H}$ and an $\hat{\epsilon} > 0$ such that $\mathcal{B}(\hat{x}, \hat{\epsilon}) \subset \mathcal{B}(\bar{x}, \epsilon)$ and $S_k \cap \mathcal{B}(\hat{x}, \hat{\epsilon}) \cap \mathcal{H} = \emptyset$. It follows from Lemma 2.4.2 that $S_k \cap \mathcal{B}(\hat{x}, \hat{\epsilon})$ is full dimensional with dimension n . For all $x \in \mathcal{B}(\hat{x}, \hat{\epsilon}) \cap S_k$ we can substitute $x_1 = F(x_2, \dots, x_n)$ into $Q_j(x) = 0$ to get $q_j(x_2, \dots, x_n) = 0$, where q_j is a polynomial function. Since $S_k \cap \mathcal{B}(\hat{x}, \hat{\epsilon})$ is full dimensional with dimension n , the set of points $\mathcal{X} = \{(x_2, \dots, x_n)^T \in \mathbb{R}^{n-1} \mid x \in S_k \cap \mathcal{B}(\hat{x}, \hat{\epsilon})\}$ contains some ball $\tilde{\mathcal{B}}$. (Note that if $(x_2, \dots, x_n)^T \in \mathcal{X}$ then it corresponds to some x with $x_1 = F(x_2, \dots, x_n)$.) Thus, since $Q_j(x) = 0$ for all $x \in S_k \cap \mathcal{B}(\hat{x}, \hat{\epsilon})$ we have that $q_j(x_2, \dots, x_n) = 0$ for

all $(x_2, \dots, x_n)^T \in \tilde{B}$. Since q_j is a polynomial, then $q_j(x_2, \dots, x_n) = 0$ for all $(x_2, \dots, x_n)^T \in \mathbb{R}^{n-1}$.

If, for a given $(x_2, \dots, x_n)^T$ we define $x_1 = F(x_2, \dots, x_n)$ so that $Q_k(x) = 0$, it then follows from the definition of $q_j(x_2, \dots, x_n)$ that we also have $Q_j(x) = 0$. Therefore, $S_k \subset S_j$. In an analogous manner, we can show that $S_j \subset S_k$ implying that $S_j = S_k$. \square

Theorem 2.4.2. Suppose that there are no pseudo quadratic constraints, no redundant constraints, and no implicit or explicit equality constraints in both (2.1.1) and (2.1.2). In what follows, the overset “-” represents items corresponding to the representation in (2.1.2). Let $\tilde{x} \in \text{int}(\mathcal{F}_j)$ for some $j \in I_N$, and let $\epsilon > 0$ be such that

$$B(\tilde{x}, \epsilon) \cap \mathcal{F}_j \subset \mathcal{F}_j^*,$$

where \mathcal{F}_j^* is a component of \mathcal{F}_j . Then there exists a unique index $k \in I_N$ such that

$$B(\tilde{x}, \epsilon) \cap \mathcal{F}_j \subset \bar{\mathcal{F}}_k^*,$$

where $\bar{\mathcal{F}}_k^*$ is a connected component of $\bar{\mathcal{F}}_k$, the face corresponding to constraint k in the representation (2.1.2).

Proof. Suppose not. Since $\tilde{x} \in \text{int}(\mathcal{F}_j)$, we have $\tilde{x} \in \partial\mathcal{R}$ and there must exist a subset \hat{I} of I_N such that

$$B(\tilde{x}, \epsilon) \cap \mathcal{F}_j \not\subset \bar{\mathcal{F}}_i^*, \quad i \in \hat{I}, \quad (2.4.3)$$

and

$$B(\tilde{x}, \epsilon) \cap \mathcal{F}_j \subset \bigcup_{i \in \hat{I}} \bar{\mathcal{F}}_i^*, \quad (2.4.4)$$

where for each $i \in \hat{I}$, $\bar{\mathcal{F}}_i^*$ is a component of $\bar{\mathcal{F}}_i$.

It follows from (2.4.3), (2.4.4), the fact that $\tilde{x} \in \partial\mathcal{R}$, and the fact that (2.1.2) contains no redundant constraints that the system obtained by adding the quadratic constraint $x \in B(\tilde{x}, \epsilon)$ to the constraint set given by \hat{I} contains no redundant

constraints. Since each constraint $i \in \hat{I}$ is necessary, it follows that there exist balls $\mathcal{B}(\bar{y}_i, \epsilon_i)$, $i \in \hat{I}$, such that

$$\mathcal{B}(\bar{y}_i, \epsilon_i) \subset \bar{\mathcal{R}}_i \setminus \mathcal{R} \quad \text{and} \quad \mathcal{B}(\bar{y}_i, \epsilon_i) \subset \mathcal{B}(\bar{x}, \epsilon)$$

for all $i \in \hat{I}$. Let $y \in \text{int}(\mathcal{R}) \cap \mathcal{B}(\bar{x}, \epsilon)$ and C_i , $i \in \hat{I}$, to be the convex hull of the set $\{y\} \cup \mathcal{B}(\bar{y}_i, \epsilon_i)$. Now define, for each $i \in \hat{I}$, the sets

$$V_i = C_i \cap \bar{\mathcal{F}}_i^* \subset \mathcal{B}(\bar{x}, \epsilon) \cap \mathcal{F}_j.$$

(see Figure 2.5). Note that each set V_i contains a set $\bar{\mathcal{F}}_i \cap \mathcal{B}(\bar{x}_i, \epsilon_i)$ for some ϵ_i sufficiently small, where \bar{x}_i is the point of intersection of $\bar{\mathcal{F}}_i^*$ with the line passing through y and \bar{y}_i . It now follows from Theorem 2.4.1 that $\bar{S}_i = \bar{S}_j$, $\forall i, j \in \hat{I}$, or that $\hat{I} = \{k\}$ for some $k \in \bar{I}_N$. Thus,

$$\mathcal{B}(\bar{x}, \epsilon) \cap \mathcal{F}_j \subset \bar{\mathcal{F}}_k^*.$$

□

2.5. Minimal Representation Theorem.

Theorem 2.5.1. The system (2.1.1) is a minimal representation of \mathcal{R} if and only if it contains no redundant constraints, no pseudo-quadratic constraints and no implicit equalities. □

Proof. We first prove that the conditions are necessary. It is straightforward to see that (2.1.1) must contain no redundant or pseudo-quadratic constraints. Suppose that (2.1.1) contains implicit equalities. Corollary 2.2.1 and the fact that there are no pseudo-quadratic constraints implies that the implicit equalities must be linear. It then follows from Lemmas 2.2.4, 2.2.5, and 2.2.6 that after the implicit equalities are replaced by explicit equalities there must then be at least one redundant constraint. This contradicts (2.1.1) being minimal and it follows that (2.1.1) contains no implicit equalities.

We now prove the sufficient conditions. Since there are no implicit equalities it follows that $\text{int}(\mathcal{R}) \neq \emptyset$ and that $\text{aff}(\mathcal{R}) = \{x \mid Ax = b\}$. Since there are no redundant constraints then the set of constraints $Ax = b$ is a minimal representation of $\text{aff}(\mathcal{R})$. Thus, for the remainder of this proof we can assume that $r = 0$. Let the system (2.1.2) be a minimal representation of \mathcal{R} . (In the proof, the “bar” indicates that the face, surface, etc., corresponds to system (2.1.2).)

Let I^j be the index set of all components of face \mathcal{F}_j , $j \in I_N$ from system (2.1.1), and let $I^{\bar{j}}$ be the index set of all components of face $\mathcal{F}_{\bar{j}}$, $\bar{j} \in I_{\bar{N}}$ from system (2.1.2). Let S_j and $S_{\bar{j}}$ be the corresponding surface sets. Since (2.1.2) is a minimal representation it follows from the first part of this proof that it contains no pseudo-quadratic constraints, no redundant constraints and no implicit equalities. Since (2.1.1) also contains no pseudo-quadratic constraints, no redundant constraints and no implicit equalities it follows from Lemma 2.4.3 and Theorem 2.4.2 that for each pair of indices $i \in I^j$, $j \in I_N$ there exists unique indices $\bar{i} \in I^{\bar{j}}$ and $\bar{j} \in I_{\bar{N}}$ such that there is a correspondence between the component \mathcal{F}_j^i and the component $\mathcal{F}_{\bar{j}}^{\bar{i}}$. That is, there exists an $\hat{x} \in \text{int}(\mathcal{F}_j^i)$ and an $\epsilon > 0$ such that $\mathcal{B}(\hat{x}, \epsilon) \cap \mathcal{F}_j^i = \mathcal{B}(\hat{x}, \epsilon) \cap \mathcal{F}_{\bar{j}}^{\bar{i}}$. From Theorem 2.4.1 it now follows that $S_j = S_{\bar{j}}$. Suppose that there are indices $s \neq j$ and $t \in I^s$ such that \mathcal{F}_s^t also corresponds to some component, say $\mathcal{F}_{\bar{j}}^{\bar{i}}$ of $\mathcal{F}_{\bar{j}}$. From Theorem 2.4.1 we get that $S_s = S_{\bar{j}}$. Thus, $S_s = S_j$ which, since there are no redundant constraints and, therefore, no duplicate constraints, contradicts $s \neq j$. Thus, $|I_N| \leq |I_{\bar{N}}|$. In a similar way we can show that $|I_{\bar{N}}| \leq |I_N|$ which implies that $|I_N| = |I_{\bar{N}}|$. Since there are no implicit equalities in either (2.1.1) or (2.1.2) it follows from Lemma 2.4.3 that all faces are full dimensional. This, in addition to the fact that there are no pseudo-quadratic constraints in either (2.1.1) or (2.1.2), allows us to use Lemma 2.4.1 to conclude that linear constraints in (2.1.1) correspond to linear constraints in (2.1.2). Conversely, we can also show that linear constraints in (2.1.2) correspond to linear constraints in (2.1.1). Thus, $|I_Q| = |I_{\bar{Q}}|$,

that is, that the number of quadratic constraints is the same in both sets (2.1.1) and (2.1.2), and (2.1.1) is a minimal representation of \mathcal{R} .

□

2.6. Identification of a Minimal Representation.

The first step in the determination of a minimal representation is to identify and remove the implicit equality constraints. We suppose that we are given a point $\hat{x} \in \mathcal{R}$. We then apply Algorithm A. In each iteration of the Algorithm we must identify the implicit equalities in a system of linear constraints. Suppose that $i \in \hat{J}_k$ and note that constraint i is an implicit equality constraint in the system

$$\{g_j^T x \leq \beta_j, j \in \hat{J}_k; B_j x = B_j \hat{x}, a_j^T x = a_j^T \hat{x}, j \in \hat{J}_0 \setminus \hat{J}_k\}$$

if and only if the LP

$$\min\{g_i^T s | g_j^T s \leq 0, j \in \hat{J}_k \setminus \{i\}; B_j s = 0, a_j^T s = 0, j \in \hat{J}_0 \setminus \hat{J}_k\}$$

is bounded from below. Suppose that the linear equality constraints force s to be in an $n - \hat{r}$ dimensional subspace. If $|\hat{J}_k|$, the cardinality of \hat{J}_k , is not greater than $(n - \hat{r} + 1)$, then these LP's can be solved in $O((n - \hat{r})^3)$ operations. Thus, the total cost is $O((n - \hat{r})^3)$ operations. If $|\hat{J}_k|$ is greater than $(n - \hat{r} + 1)$, we suggest using the techniques in [5] which were developed to solve this type of system of LP problems.

We note that Algorithm C replaces the implicit equalities with linear equalities as they are identified. However, after the identification is complete, the set of equality constraints must be reduced by eliminating the redundant equalities using Gaussian elimination. The result is a set of equalities that form a minimal representation of $\text{aff}(\mathcal{R})$. Suppose that the final system of equalities is given by $\bar{A}x = \bar{b}$, where \bar{A} is a $(\bar{r} \times n)$ matrix, i.e., we take (2.1.2) to be a minimal representation.

We now restrict all computations and comments to $\text{aff}(\mathcal{R})$. In practical terms, this implies that all quantities have been projected onto the null space of \bar{A} .

The next step is to identify all pseudo-quadratic constraints. For each of the remaining quadratic constraints we must determine if the (projected) Hessian matrix is rank one. This takes $O((n - \bar{r})^2)$ operations for each constraint. If the matrix is rank one, we must then determine if (projected) coefficients of the linear term is in the range space of the (projected) Hessian. This takes an additional $O((n - \bar{r}))$ operations for each constraint. If there are q quadratic constraints, this step takes $O(q(n - \bar{r})^2)$ operations. Each pseudo-quadratic constraint is then replaced by two linear inequalities (cf. the proof of Theorem 2.3.1).

The third and final step is to identify and remove all redundant constraints. This step is the most problematic. Since we have quadratic constraints, methods analogous to the deterministic methods given, for example, in [54], would require the solution of non convex programmes. The only realistic alternative is the use of a probabilistic method, e.g. the hypersphere directions (HD) method which was first given in [3] as the “PREDUCE” algorithm. (In fact, the motivation for the development of PREDUCE algorithm was the need to solve a large QCQP problem.) The HD method generates a sequence of feasible line segments whose end points hit the boundary of \mathcal{R} . It follows from Theorem 2.4.1 that, with probability one, each end point identifies a necessary constraint. After a finite number of iterations, all constraints that have not been identified are assumed to be redundant. As with all probabilistic methods, there is the possibility of error. In this case, it is possible that a necessary constraint was not identified and was then classified as redundant. We remark that each iteration of the HD method requires $O((m + q)(n - \bar{r})^2)$ operations, where we assume that there are $(m + q)$ inequality constraints. It follows from the results in [5] that, in the best case, i.e., in the case when all the constraints have the same probability of being detected in a single iteration of the one hit method, the expected minimal number of iterations required to detect all the necessary inequalities is bounded from above by $O(\ell \ln \ell)$, where ℓ is the actual number of

necessary constraints. Since ℓ is some fraction of $(m + q)$, the minimum expected total cost for detection is $O((m + q)^2 \ln(m + q)(n - \bar{r})^2)$.

2.7. Summary.

While the results of this chapter are mainly a theoretical contribution, we were also motivated by the potential for computational savings.

What are some of the benefits, or potential benefits, of a minimal representation theory? For one, it is reasonable to expect that a QCQP solution algorithm applied to a minimal representation would be more efficient than the same algorithm applied to any other representation. Of course, this savings has to be offset against the cost of obtaining the representation. For another, it is shown in Chapter 4 that the ability to identify and remove implicit equalities provides a new approach to the application of a method of centers to QCQP problems having “faulty feasible regions”.

Let us consider, in more detail, the potential for computational savings. Suppose that the method of centers proposed by Mehrotra and Sun [34] is being used to solve a QCQP problem. Since the algorithm needs an (relative) interior point, we assume that either there were no implicit equalities, or that the implicit equalities were identified and removed. That is, we assume that the number of linear equality constraints is \bar{r} . The Mehrotra and Sun algorithm is guaranteed to converge to the required accuracy in $O(\sqrt{m + q} L)$ iterations, where L is the bit length of the problem data. Let us look at the cost per iteration. Each iteration of the Algorithm requires the calculation of the Hessian matrix of the potential function. The terms in the Hessian matrix which correspond to linear functions require $O((n - \tau)^2)$ fewer operations than the terms which correspond to quadratic constraints. From Section 2.6, we saw that the cost of identifying all the pseudo-quadratic constraints was $O((m + q)(n - \bar{r})^2)$. Thus, if the algorithm takes $(m + q)$ iterations to converge, and if there was at least one pseudo-quadratic constraint, the effort taken to identify

the pseudo-quadratic constraints should be worthwhile.

Unfortunately, because of the high cost of an HD iteration; and because of the large expected number of iterations required to identify the necessary constraints, it appears unlikely that the redundancy detection step will be worthwhile. However, there is a potential for the use of a partial redundancy detection step to try and identify the more important constraints, that is, the constraints active at an optimal solution. We do not yet have any computational results to report.

Since pseudo-quadratic constraints show the greatest potential for computational savings, it is natural to question their occurrence in practical problems. In Hock and Schittkowski [24], problems numbered 14 and 108 have easily identified pseudo-quadratic constraints. In problem 14, the two constraints $x - 2y = -1$ and $.25x^2 + y^2 - 1 \leq 0$ combine to give the two linear constraints $.5 \leq y \leq 1$. In problem 108 the constraint $x_9^2 - 1 \leq 0$ can be replaced with $-1 \leq x_9 \leq 1$. We can also see that our characterization of pseudo-quadratic constraints depends upon convexity. In 108, we note that since $x_9 \geq 0$, the nonconvex quadratic constraints $x_3x_9 \geq 0$ and $-x_5x_9 \geq 0$ can be replaced with the linear constraints $x_3 \geq 0$ and $x_5 \leq 0$, respectively. This indicates the potential impact of pseudo-quadratic constraints in the area of nonconvex programming.

In summary, we provided a definition for a minimal representation of quadratically constrained convex feasible regions, and proved that a representation is minimal if and only if it contains no redundant constraints, no implicit equalities, and no pseudo-quadratic constraints. We also outlined the steps of a procedure that could be used to determine a minimal representation. We concluded that there are potential benefits in the determination of a minimal representation before the application of interior point methods for the solution of QCQP problems.

CHAPTER 3

BOUNDEDNESS

3.1. Introduction.

In this chapter we present an algorithm which can be used to determine whether or not a convex quadratic objective function is bounded from below over a feasible region defined by convex quadratic constraints. We say that QCQP, as given in (1.2.1) - (1.2.2), is unbounded if for any number $M \in \mathbb{R}$ there exists an $\hat{x} \in \mathcal{R}$ with $Q_0(\hat{x}) < M$. Each iteration of our algorithm requires the identification of implicit equality constraints in a system of homogeneous linear inequality and equality constraints; and the identification can be completed by using the technique in Section 1.2.2. The algorithm has the advantage of providing a mechanism to reduce both the number of constraints and the dimension of the problem. As discussed in Section 1.1, QCQP's can be unbounded even when no unbounded ray exists. This is demonstrated by the following example.

Example 3.1. The objective function is $Q_0(x) = -x_1$ and the constraints are $x_1^2 - x_2 \leq 0$, $-x_1 \leq 0$ and $-x_2 \leq 0$ (see Figure 3.4.1). Clearly, for every real number M , there is a feasible point (x_1, x_2) with $Q_0(x) < M$. Thus, $Q_0(x)$ is unbounded from below over \mathcal{R} . We will now see there is no feasible half-line along which $Q_0(x)$ is unbounded. The feasible region is unbounded along the half-line given by $(x_1, x_2) = t(0, 1)$, $t \geq 0$, but the objective is constant along that half-line. However, for every $\alpha > 0$, the feasible region, and hence the objective function, is bounded along the line $x_2 - \alpha x_1 = 0$ as it intersects the parabola at the two points $(0, 0)^T$ and $(\alpha, \alpha^2)^T$. \square

To date, methods for detecting unboundedness, have, in fact, only dealt with unbounded rays. For example, the methods in [7] which are concerned with unboundedness of both convex and concave QCQP problems, only consider the problem of

determining unbounded rays.

This chapter, which can be considered an extension of [7], presents a set of necessary and sufficient conditions for the unboundedness of a convex QCQP problem. The necessary and sufficient conditions are in the form of a novel algorithm, which, at each iteration, requires the identification of implicit equality constraints in a homogeneous linear system. Since implicit equalities can be detected with the solution of linear programmes, and since the algorithm requires at most $\min\{m, n\}$ iterations, the algorithm has the polynomial time property.

The Algorithm is used in Chapter 4 for solving QCQP problems with unbounded regions using interior point methods. Thus, our approach has other applications than detecting boundedness. In fact, we will show that each iteration of the algorithm has the added advantage of reducing both the size of the constraint set, and the dimension of the problem, by at least one.

The following two Theorems from [7] are restated, without proof, as they are important to our thesis.

Theorem 3.1.1. The region \mathcal{R} is unbounded if and only if it is nonempty and there exists a nonzero vector s satisfying the following conditions.

$$\begin{aligned} B_i s &= 0, \quad \forall i \in I \\ a_i^T s &\leq 0, \quad \forall i \in I. \end{aligned} \tag{3.1.1}$$

□

Theorem 3.1.2. Let $I_0 = I \cup \{0\}$. The function $Q_0(x)$ is unbounded from below along a ray in \mathcal{R} with direction s , if and only if the following conditions are satisfied

$$\begin{aligned} a_0^T s &< 0 \\ B_i s &= 0, \quad \forall i \in I_0 \\ a_i^T s &\leq 0, \quad \forall i \in I. \end{aligned} \tag{3.1.2}$$

□

From Theorem 3.1.1 we see that \mathcal{R} is unbounded if and only if it contains a half line $x(p, s) = \{p + \sigma s \mid \sigma \geq 0\}$, where s satisfies the conditions of the theorem. While Theorem 3.1.2 gives sufficient conditions for unboundedness, Example 3.1.1 has shown that the conditions are not necessary.

3.2. The Algorithm.

In this section we prove that the following algorithm, which terminates after at most $\min\{m, n\}$ steps, indicates whether or not $Q_0(x)$ is unbounded from below over \mathcal{R} .

Algorithm D: A Method to Determine Boundedness of the QCQP Problem

Set $k = 0$, $J_0 = I_0$

repeat

Find the set J_{k+1} of all implicit equalities
in the system $\{a_j^T s \leq 0, B_j s = 0, |j \in J_k\}$

(3.2.1)

If $0 \in J_{k+1}$ then

If $J_{k+1} \neq J_k$ then

$k := k + 1$

else stop := true { QCQP is bounded }

else stop := true { QCQP is unbounded }

until

stop

□

In order to determine the set J_k we can use the method discussed in Section 1.2.2. Clearly priority should be given to the determination of whether or not $a_0^T s \leq 0$ is an implicit equality. To continue, we need the following notation. For each index set J_k we define the regions

$$\mathcal{R}(J_k) = \{x \in \mathbb{R}^n \mid Q_i(x) \leq b_i, i \in J_k \setminus \{0\}\},$$

and

$$\mathcal{R}_z(J_k) = \{x \in \mathcal{R}(J_k) \mid Q_0(x) \leq z\},$$

where z is a fixed scalar. The matrix whose first columns are the vectors a_i , $i \in J_k$, and whose remaining columns are the columns of the matrices B_i , $i \in J_k$, is denoted

by $A(J_k)$. For example,

$$A(J_0) = [a_0, a_1, \dots, a_m, B_0, B_1, \dots, B_m].$$

The column space of $A(J_k)$ is denoted by $C(A(J_k))$, and the orthogonal complement of $C(A(J_k))$ is the null space of $A^T(J_k)$ which is denoted by $N(A^T(J_k))$. We note that if s satisfies (3.2.1), and if all constraints in J_k are implicit equalities, then $s \in N(A^T(J_k))$. Finally, for any $p \in \mathcal{R}_z(J_k)$ we define the affine space

$$\mathcal{R}_z(J_k, p) = \{p + s \mid s \in C(A(J_k))\}.$$

Lemma 3.2.1. If Algorithm D terminates with the message that $J_{k+1} = J_k$, then $\mathcal{R}_z(J_k) \cap \mathcal{R}_z(J_k, p)$ is bounded for all $p \in \mathcal{R}_z(J_k)$.

Proof. We use the method of contradiction. Suppose that $\mathcal{R}_z(J_k) \cap \mathcal{R}_z(J_k, p)$ is unbounded. This implies that $\mathcal{R}_z(J_k)$ is unbounded in the affine space $\mathcal{R}_z(J_k, p)$, so that there exists a non-zero $s \in C(A(J_k))$ such that

$$x(p, s) \subset \mathcal{R}_z(J_k) \cap \mathcal{R}_z(J_k, p).$$

Furthermore, it follows from Theorem 3.1.1 that $B_i s = 0$ and $a_i^T s \leq 0$ for all $i \in J_k$. Now termination with $J_{k+1} = J_k$, implies that $s \in N(A^T(J_k))$, which is a contradiction since $s \neq 0$ and $s \in C(A(J_k))$. Thus, $\mathcal{R}_z(J_k) \cap \mathcal{R}_z(J_k, p)$ is bounded. \square

Lemma 3.2.2. If Algorithm D terminates with the message that $J_{k+1} = J_k$, then $Q_0(x)$ is bounded from below over $\mathcal{R}(J_k)$.

Proof. For any $p \in \mathcal{R}_z(J_k)$ it follows from Lemma 3.2.1 that $\mathcal{R}_z(J_k) \cap \mathcal{R}_z(J_k, p)$ is bounded, which, in turn, implies that $Q_0(x)$ is bounded from below over $\mathcal{R}_z(J_k) \cap \mathcal{R}_z(J_k, p)$. In order to conclude that $Q_0(x)$ is bounded from below over $\mathcal{R}_z(J_k)$, and, hence, over $\mathcal{R}(J_k)$, it remains to show that for distinct $p_1 \in \mathcal{R}_z(J_k)$ and $p_2 \in \mathcal{R}_z(J_k)$ we have

$$\begin{aligned} & \min\{Q_0(x) \mid x \in \mathcal{R}_z(J_k) \cap \mathcal{R}_z(J_k, p_1)\} \\ = & \min\{Q_0(x) \mid x \in \mathcal{R}_z(J_k) \cap \mathcal{R}_z(J_k, p_2)\} \end{aligned}$$

Let W be a matrix whose columns form an orthonormal basis for $C(A(J_k))$. For $i = 1, 2$ we have $p_i = p_i^C + p_i^N$ where $p_i^C \in C(A(J_k))$ and $p_i^N \in N(A^T(J_k))$. Thus, for $i = 1, 2$, any $x \in \mathcal{R}_z(J_k, p_i)$ can be expressed as $x = p_i^N + W\xi$ for some vector ξ . Using this change of variables and the fact that $p_i^N \in N(A^T(J_k))$ we get

$$\begin{aligned} & \min\{Q_0(x) \mid x \in \mathcal{R}_z(J_k) \cap \mathcal{R}_z(J_k, p_i)\} \\ &= \min\{Q_0(p_i^N + W\xi) \mid Q_j(p_i^N + W\xi) \leq b_j, j \in J_k \setminus \{0\}\} \\ &= \min\{Q_0(W\xi) \mid Q_j(W\xi) \leq b_j, j \in J_k \setminus \{0\}\}. \end{aligned} \quad (3.2.2)$$

The result follows as the last term is independent of p_i . \square

Lemma 3.2.3. If $Q_0(x)$ is bounded from below over $\mathcal{R}(J_k)$, then $Q_0(x)$ is bounded from below over $\mathcal{R}(J_\ell)$ for $\ell > k$.

Proof. It is sufficient to show that if $Q_0(x)$ is bounded from below over $\mathcal{R}(J_k)$, then it is bounded from below over $\mathcal{R}(J_{k+1})$. It follows from Algorithm D that for each $i \in J_k \setminus J_{k+1}$ the constraint " $a_i^T s \leq 0$ " is not an implicit equality in (3.2.1), and for each $i \in J_{k+1}$ the constraint " $a_i^T s \leq 0$ " is an implicit equality in (3.2.1). Thus, for each $i \in J_k \setminus J_{k+1}$ there exists an s_i satisfying (3.2.1) with $a_i^T s_i < 0$ and with $a_j^T s_i = 0, \forall j \in J_{k+1}$. Furthermore, since $Q_0(x)$ is bounded from below over $\mathcal{R}(J_k)$ it follows from Theorem 3.1.2 that $a_0^T s_i = 0$ for each $i \in J_k \setminus J_{k+1}$. Let

$$s^* = \sum_{i \in J_k \setminus J_{k+1}} s_i.$$

We have that s^* satisfies (3.2.1) with $a_i^T s^* < 0, i \in J_k \setminus J_{k+1}, a_i^T s^* = 0, i \in J_{k+1}$, and $a_0^T s^* = 0$. Let

$$x^* = \operatorname{argmin}\{Q_0(x) \mid x \in \mathcal{R}(J_k)\}.$$

Clearly, for all $\sigma > 0$ we have $Q_0(x^* + \sigma s^*) = Q_0(x^*)$, $Q_i(x^* + \sigma s^*) = Q_i(x^*) \leq b_i, i \in J_{k+1}$, and $Q_i(x^* + \sigma s^*) < Q_i(x^*) \leq b_i, i \in J_k \setminus J_{k+1}$. Since

$$Q_0(x^* + \sigma s^*) = \min\{Q_0(x) \mid x \in \mathcal{R}(J_k)\}, \sigma > 0 \quad (3.2.3)$$

and since the constraints with $i \in J_k \setminus J_{k+1}$ are inactive at $x^* + \sigma s^*$, it follows that

$$Q_0(x^* + \sigma s^*) = \min\{Q_0(x) \mid x \in \mathcal{R}(J_{k+1})\}, \quad \forall \sigma \geq 0. \quad (3.2.4)$$

That is, $Q_0(x)$ is bounded from below over $\mathcal{R}(J_{k+1})$. \square

Equations (3.2.2), (3.2.3) and (3.2.4) have a significance well beyond the proofs of Lemmas 3.2.2 and 3.2.3. In the next Corollary, we see their implication that the algorithm can be used to reduce the number of variables and constraints in the QCQP problem.

Corollary 3.2.1. If Algorithm D terminates with the message that $J_{k+1} = J_k$, then a solution to the QCQP can be constructed from a solution to $\min\{Q_0(W\xi) \mid Q_j(W\xi) \leq b_j, j \in J_k \setminus \{0\}\}$, where $|J_k| \leq m - k$ and $\xi \in \mathbb{R}^N$, with $N \leq n - k$.

Proof. From (3.2.2) we can have that

$$\begin{aligned} & \min\{Q_0(W\xi) \mid Q_j(W\xi) \leq b_j, j \in J_k \setminus \{0\}\} \\ &= \min\{Q_0(x) \mid x \in \mathcal{R}_z(J_k)\} \\ &= \min\{Q_0(x) \mid x \in \mathcal{R}(J_k)\}. \end{aligned}$$

Let $\xi^* = \operatorname{argmin}\{Q_0(W\xi) \mid Q_j(W\xi) \leq b_j, j \in J_k \setminus \{0\}\}$ and set $x_k^* = W\xi^*$, so that $x_k^* = \operatorname{argmin}\{Q_0(x) \mid x \in \mathcal{R}(J_k)\}$. From the proof of Lemma 3.2.3, and, in particular, from (3.2.3) and (3.2.4), we see that there is a scalar σ_k and a search direction s_k^* such that $x_k^* + \sigma_k s_k^* \in \mathcal{R}(J_{k-1})$, and $Q_0(x_k^* + \sigma_k s_k^*) = \min\{Q_0(x) \mid x \in \mathcal{R}(J_{k-1})\}$. Recurring on this step, we end with $x^* = x_1^* + \sigma_1 s_1^* \in \mathcal{R}(J_0)$, and $Q_0(x^*) = \min\{Q_0(x) \mid x \in \mathcal{R}(J_0)\}$.

Since every non terminating iteration of Algorithm D determines J_{k+1} , such that $J_{k+1} \subset J_k$, it follows directly that $|J_k| \leq m - k$. Similarly, to conclude that $N \leq n - k$, we need to show that $\operatorname{rank}(C(A(J_{k+1}))) < \operatorname{rank}(C(A(J_k)))$. Suppose that $\ell \in J_k \setminus J_{k+1}$, so that $a_\ell^T s \leq 0$ is not an implicit equality in (3.2.1). Thus, there exists a vector s with $a_\ell^T s < 0$, and $B_j s = 0$ and $a_j^T s = 0$ for $j \in J_{k+1}$. Thus, $a_\ell \notin C(A(J_{k+1}))$, and the result follows. \square

Corollary 3.2.1 is demonstrated in Example 3.2.3, which appears later in this chapter.

Lemma 3.2.4. If Algorithm D terminates with the message that $0 \notin J_{k+1}$, then $Q_0(x)$ is unbounded from below over \mathcal{R} .

Proof. It follows from Theorem 3.1.2 that termination with $0 \notin J_{k+1}$, implies that $Q_0(x)$ is unbounded from below over $\mathcal{R}(J_k)$. The contrapositive of Lemma 3.2.3 implies that $Q_0(x)$ is unbounded from below over $\mathcal{R}(J_0) = \mathcal{R}$. \square

Lemma 3.2.5. If $Q_0(x)$ is unbounded from below over \mathcal{R} , then Algorithm D terminates with the message that $0 \notin J_{k+1}$.

Proof. We prove the contrapositive. Suppose that Algorithm D terminates with $J_{k+1} = J_k$. Lemma 3.2.2 then implies that $Q_0(x)$ is bounded from below over $\mathcal{R}(J_k)$. Since the constraints defining $\mathcal{R}(J_k)$ are a subset of those defining $\mathcal{R}(J_0) = \mathcal{R}$ it follows that $Q_0(x)$ is also bounded from below over \mathcal{R} . \square

The next theorem summarizes the above lemmas, that is, it summarizes the properties of the Algorithm.

Theorem 3.2.1. Algorithm D terminates with the message that $0 \notin J_{k+1}$, if $Q_0(x)$ is unbounded from below over \mathcal{R} , and terminates with the message that $J_{k+1} = J_k$, if $Q_0(x)$ is bounded from below over \mathcal{R} .

Proof. Clearly, the algorithm must terminate with either message that $J_{k+1} = J_k$ or $0 \notin J_{k+1}$. From Corollary 3.2.1, we see that termination is after at most $\min\{m, n\}$ iterations. Combining Lemmas 3.2.4 and 3.2.5 we get that termination with the message that $0 \notin J_{k+1}$ takes place if and only if $Q_0(x)$ is unbounded from below over \mathcal{R} . \square

It is worth noting that by taking $k = 0$ in Lemmas 3.2.4 and 3.2.5, we get the following corollaries, which give sufficient conditions for $Q_0(x)$ to be bounded and unbounded, respectively, over \mathcal{R} .

Corollary 3.2.2. If all inequality constraints in

$$a_i^T s \leq 0, \quad i \in J_0,$$

$$B_i s = 0, \quad i \in J_0,$$

are implicit equality constraints, then $Q_0(x)$ is bounded from below over \mathcal{R} .

Proof. Take $k = 0$ in Lemma 3.2.2. \square

Corollary 3.2.3. If there is a solution to the system

$$a_0^T s < 0,$$

$$a_i^T s \leq 0, \quad i \in J_0 \setminus \{0\},$$

$$B_i s = 0, \quad i \in J_0,$$

then $Q_0(x)$ is unbounded from below over \mathcal{R} .

Proof. Take $k = 0$ in Lemma 3.2.4. \square

3.3. Examples.

To demonstrate the effectiveness of the algorithm we present two examples. In Example 3.2.1 we apply the algorithm to the QCQP in Example 3.1.1. Recall that for this example, there was no ray along which the QCQP was unbounded. In Example 3.2.2, we apply the algorithm to a bounded QCQP with three variables. This second example first appeared in [9].

Example 3.2.1. We apply Algorithm B to the problem of Example 3.1.1. We have $a_0 = a_2 = (-1, 0)^T$, $a_1 = a_3 = (0, -1)^T$, $B_0 = B_2 = B_3 = 0$, and

$$B_1 = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}.$$

Let $s = (s_1, s_2)^T$. We begin the algorithm with $k = 0$ and $J_0 = \{0, 1, 2, 3\}$. In the next step we see that $B_i s = 0$, $i \in J_0$, implies that $s_1 = 0$, and that $a_0^T s = 0$. Thus, $a_0^T s \leq 0$ is an implicit equality and we proceed with the algorithm. Using the information from the first step of the algorithm we see that $a_0^T s \leq 0$ and $a_2^T s \leq 0$ are implicit equalities, while $a_1^T s \leq 0$ and $a_3^T s \leq 0$ reduce to $s_2 \geq 0$. Thus,

$J_1 = J_0 \setminus \{1, 3\} = \{0, 2\}$. We return to the step checking whether or not $0 \notin J_{k+1}$. We have that $B_i s = 0$, $i \in J_1$ gives no restriction on s . From $a_0^T s \leq 0$ and $a_2^T s \leq 0$ we get $s_1 \geq 0$. Thus, $a_0^T s \leq 0$ is not an implicit equality and we conclude that $Q_0(x)$ is unbounded from below over \mathcal{R} . \square

Example 3.2.2. We apply Algorithm D to the QCQP with $n=m=3$, $B = B_0 = 0$, $B_2 = 0$, $B_3 = 0$,

$$B_1 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

$a_0^T = (-1, 1, 0)$, $a_1^T = (1, 1, -1)$, $a_2^T = (1, 1, 1)$, $a_3^T = (1, -1, 0)$, $b_1 = -4$, and with $b_2 = b_3 = 10$. We start the algorithm with $k = 0$ and $J_0 = \{0, 1, 2, 3\}$.

Consider the system (3.2.1) with $k = 0$. Let $s = (s_1, s_2, s_3)^T$. From $B_1 s = 0$ we get $s_3 = -s_1 - s_2$. Using this fact, we quickly see that the inequalities $a_0^T s \leq 0$, $a_2^T s \leq 0$, and $a_3^T s \leq 0$ are the only implicit equalities. The vector $s_1^* = (-1, -1, 2)^T$ shows that $a_1^T s \leq 0$ is not an implicit equality. So the algorithm bypasses the step "If $0 \notin J_{k+1}$ ", and in step "if $J_{k+1} = J_k$ ", identifies $J_1 = \{0, 2, 3\}$. Now consider system (3.2.1) with $k = 1$. We easily see that $J_2 = \{0, 3\}$ gives the complete set of implicit equalities. The vector $s_2^* = (-1, -1, 1)^T$ shows that $a_2^T s \leq 0$ is not an implicit equality.

Finally, we consider (3.2.1) with $k = 2$ and get $J_3 = J_2$. We conclude that the QCQP is bounded from below. \square

Note that Algorithm D applied to Examples 3.2.1 and 3.2.2 required two iterations in order to determine whether or not QCQP problem is unbounded. In the latter Example, the vector a_2 is in the column space of the matrix B_1 . This causes the inequality $a_2^T s \leq 0$ to be an implicit equality in the system (3.2.1) in the first iteration. This inequality is not an implicit equality in the second iteration, because the equation $B_1 s = 0$ was removed from (3.2.1).

As was mentioned in Corollary 3.2.1, the algorithm has the additional feature

of being able to reduce the number of constraints and variables in bounded QCQP problems. This is demonstrated in our next example.

Example 3.2.3. We consider the QCQP given in Example 3.2.2, where Algorithm D terminated with $J_2 = \{0, 3\}$. Using the model QCQP in equation (3.2.2), with

$$W = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix},$$

where the column of W forms an orthonormal basis for the column space of

$$\begin{bmatrix} -1 & 1 \\ 1 & -1 \\ 0 & 0 \end{bmatrix}.$$

We get the QCQP

$$\min\{\sqrt{2}\xi \mid -\sqrt{2}\xi \leq 10\},$$

which has the solution $\xi^* = -5\sqrt{2}$ and $Q_0(\xi^*) = -10$.

Note that the original QCQP with three variable and three constraints, including one quadratic constraint, has been replaced by a linear programme in a single variable. From (3.2.3) and (3.2.4) we see that $Q_0(x^*) = Q_0(\xi^*) = -10$. However, we need to reconstruct the solution vector x^* . We can do this using (3.2.4). We start with

$$x_2^* = W\xi^* = \begin{bmatrix} 5 \\ -5 \\ 0 \end{bmatrix}.$$

We first find σ_2 such that $Q_2(x_2^* + \sigma_2 s_2^*) \leq b_2$, where s_2^* is given in Example 3.2.2. Solving for σ in $Q_2(x_2^* + \sigma s_2^*) \leq b_2$ yields $\sigma \geq -10$, so we take $\sigma_2 = 0$ and set $x_1^* = x_2^*$.

We now find σ_1 such that $Q_1(x_1^* + \sigma_1 s_1^*) \leq b_1$, where s_1^* is given in Example 3.2.2. Solving for σ in $Q_1(x_1^* + \sigma s_1^*) \leq b_1$ yields $\sigma \geq 1$, so we take $\sigma_1 = 1$ and set $x^* = x_1^* + s_1^* = (4, -6, 2)^T$. A quick check of the optimality conditions for the original

QCQP shows that this is indeed the solution with the first and third constraint being the active constraints. Note that there are alternate optimal solutions, which are obtained by taking $\sigma_1 \geq 1$. For $\sigma_1 > 1$, only the third constraint will be active. In fact, we actually have a two parameter family of alternate solutions with parameters σ_1 and σ_2 .

3.4. Summary.

In [7] the authors presented necessary and sufficient conditions for the existence of a feasible half line along which the objective function is unbounded. However, the existence of such a half-line is not a necessary condition for unboundedness. In this chapter, we presented necessary and sufficient conditions for unboundedness. The conditions take the form of an algorithm, the implementation of which requires the determination of implicit equality constraints in a homogeneous linear system. An added benefit of the algorithm is that each iteration reduces the number of constraints, and the dimension of the problem, by at least one. The result is that, if the QCQP is bounded, a QCQP with fewer constraints and variables can be solved to obtain the solution.

CHAPTER 4

FAULTY FEASIBLE REGIONS

4.1. Introduction.

In this chapter we are concerned with the application of the method of centers to problems having a faulty feasible region, i.e., a region that is either unbounded or not full dimensional.

Each iteration of a method of centers (see Algorithm B in Section 1.2.1.2) must determine an center of a "truncated feasible region". If the feasible region is faulty, then the truncated region may have no interior, it may have no analytic center, it may have an infinity of centers, or it may have a unique center. Thus, the difficulty caused by faulty feasible regions is that, even when the problem has a solution, the method of centers may fail.

Typically, faulty regions have been dealt with by the direct method of adding more constraints and variables, often involving a "Big M" constant. These additions adversely affect the numerical conditioning of the problem. (A notable exception is the paper by Lustig [32] in which a Big M method for linear programs is given which avoids the difficulties with numerical conditioning.) Our method is unique in that it results in a reduction in the number of constraints and, effectively, in the number of variables.

In Section 4.2 we present a new way to deal with unbounded regions. Our approach results in an effective reduction in the number of variables and a reduction in the number of constraints. It does not involve Big M constants. Section 4.3 deals with regions that are not full dimensional.

4.2. Unbounded Feasible Regions.

In this section we assume that \mathcal{R} is full dimensional, but unbounded. The first step is to determine whether or not $\mathcal{R}(z)$ is unbounded. This can be done using

slight modification of Theorem 3.1.1.

Theorem 3.1.1 with $I = I_0$, gives that the region $\mathcal{R}(z)$ is unbounded if and only if it is non-empty and if there is a vector $s \neq 0$ satisfying

$$(4.2.1) \quad \begin{aligned} B_i s &= 0, & \forall i \in I_0, \\ a_i^T s &\leq 0, & \forall i \in I_0. \end{aligned}$$

If $\mathcal{R}(z)$ is bounded, then it will have a unique center and Algorithm B will converge to an optimal solution of the QCQP. So, we will assume that $\mathcal{R}(z)$ is unbounded. The next step is to determine whether or not $\mathcal{R}(z)$ has a center.

Theorem 4.2.1. A center of the nonempty set $\mathcal{R}(z)$ exists if and only if each inequality in the system (4.2.1) is an implicit equality.

Proof. If $\mathcal{R}(z)$ is bounded then the result follows from Theorem 3.1.1 as it implies that the only solution to (4.2.1) is $s = 0$. Now suppose that $\mathcal{R}(z)$ is unbounded. We prove the forward implication by contradiction. Suppose that there are some constraints in (4.2.1) that are not implicit equalities. In particular, assume that there is a solution \bar{s} to (4.2.1) together with an index κ such that $a_\kappa^T \bar{s} < 0$. Since $\text{int}\mathcal{R}(z) \neq \emptyset$ there exists an $\bar{x} \in \text{int}\mathcal{R}(z)$. Now, for all $j \in I_0$, we have $Q_j(\bar{x} + \sigma \bar{s}) - b_j \leq Q_j(\bar{x}) - b_j < 0$, for all $\sigma \geq 0$. This implies that $(\bar{x} + \sigma \bar{s}) \in \mathcal{R}(z)$ for all $\sigma \geq 0$. Since $a_\kappa^T \bar{s} < 0$ and $B_\kappa \bar{s} = 0$ we also have that

$$\lim_{\sigma \rightarrow \infty} (b_\kappa - Q_\kappa(\bar{x} + \sigma \bar{s})) = +\infty,$$

which implies that $F(x, z)$ has no minimum over $\mathcal{R}(z)$. This contradicts the existence of an analytic center for $\mathcal{R}(z)$, and proves the forward implication.

We now prove the backwards implication. Since each inequality in (4.2.1) is an implicit equality we have that

$$(4.2.2) \quad \begin{aligned} (B_j s = 0, a_j^T s \leq 0, \forall j \in I_0) &\implies \\ (a_j^T s = 0, \forall j \in I_0). \end{aligned}$$

We will show that $\mathcal{R}(z)$ has an infinity of analytic centers. For any point $\bar{x} \in \mathcal{R}(z)$ we define the manifold

$$\mathcal{M}(\bar{x}) = \{ x \mid x = \bar{x} + p, p \in C(\nabla^2 F(x, z)) \}.$$

To establish that $\mathcal{R}(z) \cap \mathcal{M}(\bar{x})$ has an analytic center, we need only show that it is bounded. We use the method of contradiction. Suppose $\mathcal{R}(z) \cap \mathcal{M}(\bar{x})$ is unbounded. Since $\mathcal{R}(z)$ is unbounded in the linear manifold $\mathcal{M}(\bar{x})$, it follows from Theorem 3.1.1 that there exists a non zero $p \in C(\nabla^2 F(x, z))$ such that

$$\{ x \mid x = \bar{x} + tp, t \geq 0 \} \subset \mathcal{R}(z) \cap \mathcal{M}(\bar{x}),$$

where $B_j p = 0$ and $a_j^T p \leq 0$ for all $j \in I_0$. From (4.2.2) we conclude that $p \in N(\nabla^2 F(x, z))$. Since $p \neq 0$ this contradicts $p \in C(\nabla^2 F(x, z))$. Thus, $\mathcal{R}(z) \cap \mathcal{M}(\bar{x})$ is bounded, and, by consequence, contains an analytic center.

Let \bar{x}_1 and \bar{x}_2 be two distinct points in $\mathcal{R}(z)$ and let \hat{x}_1 and \hat{x}_2 be the analytic centers of $\mathcal{R}(z) \cap \mathcal{M}(\bar{x}_1)$ and $\mathcal{R}(z) \cap \mathcal{M}(\bar{x}_2)$, respectively. (It is helpful to refer to Fig. 4.1.) If we can show that $F(\hat{x}_1, z) = F(\hat{x}_2, z)$, then we can conclude that any point on line passing through \hat{x}_1 and \hat{x}_2 is an analytic center of $\mathcal{R}(z)$. From the definition of $\mathcal{M}(\bar{x})$ it follows that there exists a vector $\hat{s} \in N(\nabla^2 F(x, z))$ such that $\hat{x}_1 + \hat{s} \in \mathcal{M}(\bar{x}_2)$. Since $\hat{s} \in N(\nabla^2 F(x, z))$ it follows that $Q_j(\hat{x}_1 + \hat{s}) = Q_j(\hat{x}_1)$, for all $j \in I_0$, so that, in fact, $\hat{x}_1 + \hat{s} \in \mathcal{R}(z) \cap \mathcal{M}(\bar{x}_2)$, and $F(\hat{x}_1 + \hat{s}, z) = F(\hat{x}_1, z)$. Analogously, we also have that $F(\hat{x}_2, z) = F(\hat{x}_2 - \hat{s}, z)$, where $\hat{x}_2 - \hat{s} \in \mathcal{R}(z) \cap \mathcal{M}(\bar{x}_1)$. We will use contradiction to establish that $F(\hat{x}_1 + \hat{s}, z) = F(\hat{x}_2, z)$, which will prove the result. Suppose that $F(\hat{x}_1 + \hat{s}, z) > F(\hat{x}_2, z)$, which implies that $F(\hat{x}_2 - \hat{s}, z) < F(\hat{x}_1, z)$. This contradicts \hat{x}_1 being an analytic center of $\mathcal{R}(z) \cap \mathcal{M}(\bar{x}_1)$. Thus, $F(\hat{x}_1 + \hat{s}, z) = F(\hat{x}_2, z)$. We note that $\hat{x}_2 = \hat{x}_1 + \hat{s}$ since the analytic center of $\mathcal{R}(z) \cap \mathcal{M}(\bar{x}_2)$ is unique.

□

In order to check the conditions of Theorem 4.2.1, we need only determine whether or not any of the $(m + 1)$ LPs of the form “ $\min\{a_k^T s \mid B_j s = 0, a_j^T s \leq 0, j \in I_0\}$ ” are unbounded from below. If $a_k^T s$ is unbounded, then $a_k^T s \leq 0$ is not an implicit equality (cf. Section 1.2.2 or Telgen [53].)

Suppose that Theorem 4.2.1 determined that $\mathcal{R}(z)$ has an analytic center. It follows from the proof of Theorem 4.2.1 that $\mathcal{R}(z)$ has an infinity of analytic centers. Referring again to Fig. 4.1, we see that $\hat{x}_1 + \sigma(\hat{x}_2 - \hat{x}_1)$ is also an analytic center of $\mathcal{R}(z)$ for any value of $\sigma \in \mathbb{R}$. Also, for any $\bar{x} \in \mathcal{R}(z)$, there corresponds an analytic center of $\mathcal{R}(z)$ in the set $\mathcal{R}(z) \cap \mathcal{M}(\bar{x})$.

Even though $\mathcal{R}(z)$ has an infinity of analytic centers, we are still faced with the task of computing one of them. That is, we must still solve the Newton equations (4.1.2), which, in this case, have a singular coefficient matrix.

Let x be an arbitrary point in \mathcal{R} .

Lemma 4.2.1. If $\mathcal{R}(z)$ is unbounded but has an analytic center, then $\nabla^2 F(x, z)$ is singular.

Proof. Since $\mathcal{R}(z)$ is unbounded it follows from Theorem 3.1.1 that there is a non zero vector s satisfying (4.2.1). Since $\mathcal{R}(z)$ has an analytic center it follows from Theorem 4.2.1 that all inequalities in (4.2.1) are implicit equalities. It then follows from (4.1.3) that $s \in N(\nabla^2 F(x, z))$. Thus, $N(\nabla^2 F(x, z)) \neq \emptyset$.

□

This difficulty with singularity is easily overcome by solving the projected Newton equations

$$W^T \nabla^2 F(x, z) W \xi = -W^T \nabla F(x, z),$$

where W is an orthonormal basis for $C(\nabla^2 F(x, z))$ and $\xi = W^T p$. Note that $W^T \nabla^2 F(x, z) W$ is positive definite, and recall that $C(\nabla^2 F(x, z))$ is independent of x and z . Clearly, solving the projected Newton equations is equivalent to solving

the reduced variable QCQP given by

$$(4.2.3) \quad \min\{Q_0(W\xi) \mid Q_j(W\xi) \leq b_j, j \in I\}.$$

So, we have shown that the method of analytic centers is easily modified to handle the unboundedness of $\mathcal{R}(z)$, provided that $\mathcal{R}(z)$ has an analytic center. Furthermore, we did it in a way that effectively reduced the number of variables, rather than adding variables and/or constraints.

Now suppose that Theorem 4.2.1 determined that $\mathcal{R}(z)$ has no analytic center. We must now determine whether or not the QCQP is unbounded. This can be done with the Algorithm D, given in Section 3.2 of this thesis (we recall that this algorithm terminates after a finite number of iterations).

We suppose that Algorithm D determined that the QCQP is bounded. However, since the truncated region has no analytic center, the method of centers will fail. We will use the following result to show that we can determine, and then solve with the method of centers, an equivalent QCQP having fewer variables and constraints.

Lemma 4.2.2. Suppose that the QCQP is bounded, and that Algorithm D terminates with the index set I_k . Let QCQP_k denote the problem $\min\{Q_0(x) \mid Q_j(x) \leq b_j, j \in I_k \setminus \{0\}\}$. Any solution x^* to the QCQP is also a solution to QCQP_k.

Proof. [Follows from the proof of Corollary 3.2.1 in Chapter 3.]

□

Lemma 4.2.2 indicates that, rather than solving the QCQP, we should consider solving QCQP_k. An immediate advantage is that QCQP_k has fewer constraints. But there are more significant advantages. Since Algorithm D terminated with the QCQP being bounded, all inequalities in the system $\{a_j^T s \leq 0, B_j s = 0, |j \in I_k\}$ are implicit equalities. Furthermore, $\{0\} \in I_k$. It then follows from Theorem 4.2.1 that QCQP_k has truncated feasible regions with analytic centers. Consequently, QCQP_k can be solved with a method of centers. However, since $\mathcal{R}(z)$ is unbounded,

the truncated regions for QCQP_k will also be unbounded. Therefore, the method of centers must use the projected Newton equations corresponding to (4.2.3) with $I = I_k$ and with W modified to account for the deleted constraints.

A difficulty with this approach is that a solution to QCQP_k is not necessarily feasible for the QCQP. Fortunately, it is relatively easy to move from a solution of QCQP_k to a solution of the QCQP (see Chapter 3). Suppose that Algorithm D, terminated with the index set $I_k = I_{k+1}$. This implies that for each index $\ell \in I_{k-1} \setminus I_k$ there exists a vector s_k^ℓ with $a_\ell^T s_k^\ell < 0$, $a_j^T s_k^\ell \leq 0$, $j \in I_{k-1}$, $j \neq \ell$, and $B_j s_k^\ell = 0$, $j \in I_{k-1}$. Let x_k be a solution to QCQP_k , and define

$$(4.2.4) \quad s_k = \sum_{\ell \in I_{k-1} \setminus I_k} s_k^\ell.$$

It follows that there exists a $\sigma_k \geq 0$, which can be obtained using the usual “step size” operations, such that $x_{k-1} = x_k + \sigma_k s_k$ is feasible for QCQP_{k-1} . Clearly, $Q_0(x_k) = Q_0(x_{k-1})$, so that x_{k-1} is a solution to QCQP_{k-1} . We repeat the process until a solution to the original problem, QCQP_0 , is obtained.

Consider the following examples. In Example 4.2.1 we see the geometry, and in Example 4.2.2 we see algebraic details.

Example 4.2.1. Consider Fig. 4.2, where $\mathcal{R}(z)$ is given by two quadratic constraints along with the objective function constraint. The three constraints are indicated by their boundaries $Q_0(x) = z$, $Q_1(x) = b_1$, and $Q_2(x) = b_2$. We see that \mathcal{R} and $\mathcal{R}(z)$ are unbounded, that $\mathcal{R}(z)$ has no analytic center, but that the QCQP has a solution. Algorithm C would terminate with the index set $\hat{J} = \{0, 1\}$. We remove constraint $Q_2(x) \leq b_2$. Take \bar{x} as the initial feasible point, and, restrict the next iterates to be in the linear manifold that passes through \bar{x} and is orthogonal to $N(\nabla^2 F_{\hat{J}}(x, z))$, where $F_{\hat{J}}(x, z)$ is obtained using constraints $j \in \hat{J}$. (That is, we are effectively solving the reduced variable problem (4.2.3) with $I = \{0, 1\}$, and with W having as columns a basis for $C(\nabla^2 F_{\hat{J}}(x, z))$.) With the second constraint

removed, the method of centers will produce x_j^* as the optimal solution to the reduced problem. We now choose an $s^* \in N(\nabla^2 F_j(x, z))$ satisfying $B_2 s^* = 0$ and $a_2^T s^* < 0$. Finally, we determine a positive scalar σ^* , to get the optimal solution $x^* = x_j^* + \sigma^* s^*$ to the original problem.

□

Example 4.2.2. We apply Algorithm D to the QCQP with $n=m=3$, $B = B_0 = 0$, $B_2 = 0$, $B_3 = 0$,

$$B_1 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix},$$

$a_0^T = (-1, 1, 0)$, $a_1^T = (1, 1, -1)$, $a_2^T = (1, 1, 1)$, $a_3^T = (1, -1, 0)$, $b_1 = -4$, and with $b_2 = b_3 = 10$. We start the algorithm with $k = 0$ and $I_0 = \{0, 1, 2, 3\}$.

Consider the system (4.2.1) with $I = I_0$. Let $s = ((s)_1, (s)_2, (s)_3)^T$. From $B_1 s = 0$ we get $(s)_3 = -(s)_1 - (s)_2$. Using this fact, we quickly see that the inequalities $a_0^T s \leq 0$, $a_2^T s \leq 0$, and $a_3^T s \leq 0$ are the only implicit equalities. The vector $s_1^1 = (-1, -1, 2)^T$ shows that $a_1^T s \leq 0$ is not an implicit equality. So the algorithm identifies $I_1 = \{0, 2, 3\}$.

Now consider system (4.2.1) with $I = I_1$. The complete set of implicit equalities is given by $I_2 = \{0, 3\}$. The vector $s_2^2 = (-1, -1, 1)^T$ shows that $a_2^T s \leq 0$ is not an implicit equality.

Finally, we consider (4.2.1) with $I = I_2$ and get $I_3 = I_2$. We conclude that the QCQP is bounded from below. The matrix W has columns which form an orthonormal basis for

$$C[a_0, a_3, B_0, B_3] = C \begin{bmatrix} -1 & 1 \\ 1 & -1 \\ 0 & 0 \end{bmatrix}.$$

We can take $W = [-1/\sqrt{2}, 1/\sqrt{2}, 0]^T$, to get the reduced QCQP

$$\min\{\sqrt{2}\xi \mid -\sqrt{2}\xi \leq 10\},$$

which has the solution $\xi^* = -5\sqrt{2}$ and $Q_0(\xi^*) = -10$. (The method of centers would be used to find the solution.) Note that the original QCQP with three variables and three constraints, including one quadratic constraint, has been replaced by a linear programme in a single variable. We now reconstruct the solution vector x^* . We start with

$$x_2^* = W\xi^* = \begin{bmatrix} 5 \\ -5 \\ 0 \end{bmatrix}.$$

(We note that in practice, there is no need, nor is it desirable, to actually carry out the change of variables.) Using (4.2.4) we obtain $s_1 = (-1, -1, 2)^T$ and $s_2 = (-1, -1, 1)^T$. Solving for σ in $Q_2(x_2^* + \sigma s_2^*) \leq b_2$ yields $\sigma \geq -10$, so we take $\sigma_2 = 0$ and set $x_1^* = x_2^*$. Solving for σ in $Q_1(x_1^* + \sigma s_1^*) \leq b_1$ yields $\sigma \geq 1$, so we take $\sigma_1 = 1$ and set $x^* = x_1^* + s_1^* = (4, -6, 2)^T$. A quick check of the optimality conditions for the original QCQP shows that this is indeed the solution with the first and third constraints being the active constraints.

□

We have shown how a method of centers, together with Newton's method for approximating the centers, can be applied to all bounded QCQPs having a full dimensional region. Our approach involves the solution of subsidiary LP problems to determine a reduction in the number of constraints, and an effective reduction in the number of variables. It does not involve the addition of constraints, the addition of variables, or the use of Big M constants. Furthermore, our method identifies a k parameter family of alternate optima.

We first show that the set of vectors s_k given by (4.2.4) are linearly independent. Suppose that $s_k = \alpha_1 s_1 + \dots + \alpha_{k-1} s_{k-1}$. It follows from Algorithm D that $a_j^T s_i = 0$ and $B_j s_i = 0$ for all $j \in J_i$, $i = 1, \dots, k-1$, and that $J_{k-1} \subset J_{k-2} \subset \dots \subset J_1$. Together, these statements imply that $a_j^T s_k = 0$ and $B_j s_k = 0$ for $j \in J_{k-1}$. This contradicts $a_j^T s_k < 0$ for $j \in J_{k-1} \setminus J_k$, which follows from the definition of s_k .

Thus, the vectors are linearly independent. Now we observe that

$$x^* = x_k^* + \sum_{i=1}^k \sigma_i s_i$$

is an optimal solution to QCQP provided that the parameters σ_i , $i = 1, \dots, k$, are chosen to guarantee feasibility.

4.3. Feasible Regions That Are Not Full Dimensional: Initial Interior Point.

In this section we assume that \mathcal{R} is not full dimensional. The difficulty is that there is no interior point. Our approach presented in Chapter 2 is to first identify all linear and quadratic implicit equality constraints and replace them with explicit linear equality constraints. The result is a new, equivalent QCQP having fewer constraints, a full dimensional feasible region, and an effective reduction in the number of variables.

The key to the approach presented in Algorithm C of Chapter 2 is the fact that quadratic implicit equalities can be replaced by a system of linear equality constraints. It was also indicated in Chapter 2 that many of these linear constraints would be redundant.

Suppose that Algorithm C given in Chapter 2 terminates with the index set \hat{J}_k , and consider QCQP with the new representation of the feasible region \mathcal{R} given by

$$\{Q_j(x) \leq b_j, j \in \hat{J}_k \cup I \setminus \hat{J}_0; B_j x = B_j \bar{x}, a_j^T x = a_j^T \bar{x}, j \in \hat{J}_0 \setminus \hat{J}_k\}.$$

We must now construct an x_0 in the relative interior of \mathcal{R} . From Lemma 2.2.3 it follows that each constraint $Q_j(x) \leq b_j$, $j \in \hat{J}_k$, is not an implicit equality in (4.1.1). For each index $\ell \in \hat{J}_k$, the algorithm constructs a vector s^ℓ satisfying $(\nabla Q_\ell(\bar{x}))^T s^\ell < 0$, $(\nabla Q_j(\bar{x}))^T s^\ell \leq 0$, $j \in \hat{J}_k \setminus \{\ell\}$, $B_j s^\ell = 0$, $j \in \hat{J}_0 \setminus \hat{J}_k$, and $a_j^T s^\ell = 0$, $j \in \hat{J}_0 \setminus \hat{J}_k$. We define

$$s = \sum_{\ell \in \hat{J}_k} s^\ell,$$

and set $x_0 = \bar{x} - ts$, where $t > 0$ is a suitably chosen step size to guarantee that x_0 is in the relative interior of \mathcal{R} . The problem can now be solved by a method of analytic centers.

4.4. Faulty Representations.

While the concept of an analytic center is a geometrical one, we will see that the computed analytic center, and, therefore, the performance of a method of analytic centers, depends on the representation of the region. We assume that \mathcal{R} is both bounded and full dimensional, but that the representation given in (1.2.2) is faulty, that is, that it contains redundant and/or pseudo quadratic constraints. We will also use the term *geometric analytic center* to refer to the analytic center obtained from a minimal representation of the feasible region.

The following lemma states conditions under which the addition of constraints changes the analytic center.

Lemma 4.4.1 Let system A be given by the set of convex inequalities $\{Q_i(x) \leq b_i \mid i \in I\}$, and let system B be the set of convex inequalities $A \cup \{\hat{Q}_i(x) \leq \hat{b}_i \mid i \in \hat{I}\}$, where $\hat{I} = \{1, 2, \dots, \hat{m}\}$. Let \mathcal{R}_A and \mathcal{R}_B represent the feasible regions defined by systems A and B , respectively. Suppose that $\mathcal{R}_B \subseteq \mathcal{R}_A$. If

$$(4.4.1) \quad \mathcal{S} = \{s \in \mathbb{R}^n \mid \hat{B}_i s = 0, \hat{a}_i^T s < 0, i \in \hat{I}\} \neq \emptyset$$

then there exists no point \bar{x} that is an analytic center for both \mathcal{R}_A and \mathcal{R}_B .

Proof. If either \mathcal{R}_A or \mathcal{R}_B has no analytic center, the result is trivial. Suppose that both \mathcal{R}_A and \mathcal{R}_B have analytic centers. The proof is by contradiction. Let x_0 be an analytic center of both \mathcal{R}_A and \mathcal{R}_B . Thus, x_0 is a solution to both

$$(4.4.2) \quad \nabla F(x) = \sum_{i=1}^m \frac{B_i x + a_i}{b_i - Q_i(x)} = 0,$$

and

$$(4.4.3) \quad \sum_{i=1}^m \frac{B_i x + a_i}{b_i - Q_i(x)} + \sum_{i=1}^{\hat{m}} \frac{\hat{B}_i x + \hat{a}_i}{\hat{b}_i - \hat{Q}_i(x)} = 0$$

Subtracting (4.4.2) from (4.4.3) we get

$$\sum_{i=1}^m \frac{\hat{B}_i x_0 + \hat{a}_i}{\hat{b}_i - \hat{Q}_i(x_0)} = 0$$

which implies that

$$\sum_{i=1}^m \frac{s^T \hat{B}_i x_0 + \hat{a}_i^T s}{\hat{b}_i - \hat{Q}_i(x_0)} = 0,$$

for all $s \in \mathcal{S}$, which contradicts (4.4.1). \square

There are several interesting points to be made. The first is that if only redundant constraints are added, i.e., if $\mathcal{R}_B = \mathcal{R}_A$, then Lemma 4.4.1 implies that the computed analytic centers can differ. It should also be noted that if $\mathcal{R}_B \subset \mathcal{R}_A$, the analytic centers can still be the same for both regions.

As an example consider the system A given by the set of constraints $-e \leq x \leq e$, where $e \in \mathbb{R}^2$ is a vector of ones, and the system B given by the set of constraints in A together with k duplications of the constraint $x_1 \leq 1$. Both systems define the same feasible region, and the geometric analytic center of the region is $x = 0$. While system A yields the geometric center $x = 0$, system B yields the center $(-k/(k+2), 0)$. Notice that even a single redundant constraint can have a dramatic effect on the analytic center, e.g., when k changes from 0 to 1 the center changes from $(0, 0)$ to $(1/3, 0)$. We note that the duplication of a constraint effectively changes the “weight” assigned to that constraint in the definition of the potential function.

We now consider how redundant constraints can influence the rate of convergence. Consider system A as given above together with the system B. Letting x^* be the solution and x_j be the current estimate of the solution, we have

$$\frac{\|x^* - x_{j+1}\|}{\|x^* - x_j\|} = \frac{k}{k+1}.$$

Thus, the convergence slows as k increases. Conversely, the same example can be used to show that the convergence rate can be increased by removing redundant

constraints. In fact, a stronger statement is implied by Theorem 3.1 of [36]. Suppose that $(1 - \xi)100\%$ of the constraints in (1.1) are redundant. The rate of convergence is then given by

$$\frac{z^* - z_+}{z^* - z_c} \leq 1 - \frac{.001}{\sqrt{\xi m}},$$

where $z_+ = Q(x_+)$, $z_c = Q(x_c)$ and $z^* = Q(x^*)$.

We now consider pseudo-quadratic constraints. It has been shown in the proof of the Theorem 2.3.1 that pseudo-quadratic constraints can be replaced by two linear inequalities.

$$(4.4.4) \quad \begin{aligned} u_k^T x &\leq r_k^1 \\ u_k^T x &\geq r_k^2 \end{aligned}$$

where $B_k = u_k u_k^T$, $a_k = \alpha u_k$ and $r_k^1 = -\alpha + \sqrt{\alpha^2 + 2b_k}$ and $r_k^2 = -\alpha - \sqrt{\alpha^2 + 2b_k}$.

The formula for $\nabla F(x)$ in (4.4.2) together with the inequalities in (4.4.2) imply that if a pseudo-quadratic constraint is replaced by two linear inequalities then the analytic center will not change. However, this is not the case when the pseudo-quadratic constraint can be replaced by a single linear inequality, i.e., in the case when one of the linear constraints is redundant.

Lemma 4.4.2.2. Let

$$(4.4.5) \quad \begin{aligned} Q_i(x) &\leq b_i, & i \in I \\ \hat{Q}_i(x) &\leq \hat{b}_i, & i \in \hat{I} \end{aligned}$$

be such that the constraints in \hat{I} are the complete set of pseudo-quadratic constraints, and such that there are no implicit equality constraints in (4.4.5). Assume that $I_P \subset \hat{I}$ are indices of the pseudo-quadratic constraints which can be replaced by the single linear inequality $u_j^T x \leq r_j^1$. Then the analytic center of the system (4.4.5) and the analytic center of the system

$$(4.4.6) \quad \begin{aligned} Q_i(x) &\leq b_i, & i \in I \\ \hat{Q}_i(x) &\leq \hat{b}_i, & i \in \hat{I} \setminus I_P \\ u_j^T x &\leq r_j^1, & j \in I_P \end{aligned}$$

are the same if and only if the analytic center of (4.4.6) is an analytic center of the system

$$(4.4.7) \quad u_j^T x \geq r_j^2, \quad i \in I_P$$

Proof. Straightforward.

□

As an example, consider the following two representations of the same feasible region.

$$\left\{ \begin{array}{l} -x + y \leq 0 \\ -x - y \leq 0 \\ x^2 \leq 1 \end{array} \right\} \quad (\text{I}) \quad \text{and} \quad \left\{ \begin{array}{l} -x + y \leq 0 \\ -x - y \leq 0 \\ x \leq 1 \end{array} \right\} \quad (\text{II}).$$

The analytic center of representation I is $\left(\frac{\sqrt{5}-1}{2}, 0\right)$; while for representation II it is the point $\left(\frac{2}{3}, 0\right)$. The quadratic constraint in representation I can be replaced by the two linear constraints $x \geq -1$ and $x \leq 1$, the first of which is redundant. The difference between these analytic centers shows that the redundant linear constraint repels the analytic center.

In order to have an unambiguous notion of an analytic center, we say that the analytic center obtained from a minimal representation of the quadratically constrained region is the “geometrical analytic center”. In Chapter 2 it was proved that a minimal representation is one containing no redundant constraints, no implicit equalities, and no pseudo-quadratic constraints.

4.5. Summary.

In this chapter we presented a new method for dealing with the application of a method of analytic centers to QCQP problems having a faulty feasible region. There are two kinds of difficulties which can occur, namely, unboundedness and lack of full dimensionality.

If the initial feasible point is not in the relative interior of \mathcal{R} , then Algorithm C of Chapter 2 can be used to determine all linear and quadratic implicit equality

constraints, which can then be replaced by linear equalities (many of which will be redundant). This gives an effective reduction in the number of variables. Suppose that the initial point is in the relative interior of \mathcal{R} . We apply Algorithm D of Chapter 3 to determine if the QCQP is bounded. If it is bounded, the algorithm identifies QCQP_k whose truncated feasible regions have analytic centers. If the original QCQP has a bounded region $\mathcal{R}(z)$, then Algorithm D will stop with $J_k = J_0$. Otherwise, QCQP_k will have fewer constraints than QCQP . The method of centers can then be used to determine a solution to QCQP_k , and simple step size routines can recover a solution to the original problem. If $\mathcal{R}(z)$ is unbounded, then this last step may require the use of the projected Newton equations, giving another effective reduction in the number of variables. Finally, we highlighted the difference between the analytic center of a representation and the geometrical analytic center, and we have shown that a faulty representation of a feasible region can affect the convergence rates of the method of analytic centers.

Appendix 4A

We present some minor results that arise from this chapter.

Corollary 4B.1. If $\mathcal{R}(z)$ is unbounded, then either the QCQP is unbounded from below or the QCQP has alternate optimal solutions.

Proof. We prove the contrapositive. Let \bar{x} be the unique minimizer of the QCQP and let $\bar{z} = Q(\bar{x})$. Since \bar{x} is unique, it follows that $\mathcal{R}(\bar{z})$ is bounded. Thus, $\mathcal{R}(\bar{z})$ is bounded for all truncation levels \bar{z} .

□

Corollary 4B.2. If $\mathcal{R}(z)$ has an analytic center then the QCQP is bounded.

Proof. Suppose that $\mathcal{R}(z)$ has an analytic center. Theorem 4.2.1 implies that all constraints in (4.2.1) are implicit equalities. It then follows from Algorithm D, that the QCQP is bounded.

□

CHAPTER 5

IMPROVING THE RATE OF CONVERGENCE

5.1. Introduction.

Of particular interest to this chapter is the work by Kovacevic-Vujcic [29]. In [29], the author is concerned with improving the rate of convergence of interior point methods for linear programming. In particular, the author shows that if $\{x^k\}$ is a sequence generated by an interior point method, then an auxiliary sequence $\{\bar{x}^k\}$ can be constructed which converges superlinearly faster to the solution than the original sequence $\{x^k\}$. The author then shows how the method can be applied to Karmarkar's projective algorithm [28], and to the affine scaling algorithms of Dikin [14] and Barnes [1].

This chapter is concerned with the application of the technique in [29] to a logarithmic barrier function method (Algorithm A) and to a method of centers (Algorithm B) for the solution of the QCQP. Although the results are presented only for quadratic constraints, they are valid for twice continuously differentiable convex functions, with the proofs of the theorems unchanged. As we are concerned with different algorithms applied to a more general problem than in [19], our method requires a substantially different approach to the proofs. More particularly, we consider a convergent sequence $\{x^k\}$ generated by the application of a logarithmic barrier function method or a method of analytic centers to a smooth convex programme. We first show how to construct an auxiliary sequence $\{\bar{x}^k\}$ that converges "superlinearly faster" than the original sequence $\{x^k\}$ (Section 5.2). We then give new results which show how the auxiliary sequence can be used to develop a modified method of centers having a superlinear rate of convergence (Section 5.3). The characteristic feature of our modified method is that it *predicts* the next iterate by using a *long step* along the line passing through the two points x^k and \bar{x}^k , from the

original and auxiliary sequences, respectively. In order to *correct* the prediction, a sequence of Newton steps is used to obtain the next point x^{k+1} on the optimal trajectory. In this sense we can classify our method as a predictor corrector method. In Section 5.4, we present a few numerical examples to demonstrate the potential of our modified method of centers as a practical solution algorithm for solving QCQP problems. Our limited computational experience suggests that our method could compete successfully with other interior point methods. Conclusions are in Section 5.5.

Without loss of generality, we consider the problem (1.2.1) - (1.2.2) with $B_0 = 0$, and with $a_0 = c$, that is, the problem

$$\begin{aligned} & \text{minimize} && c^T x \\ & \text{subject to} && x \in \mathcal{R}, \end{aligned} \tag{5.1.1}$$

Our assumption that the objective function is linear does not cause a loss of generality. Any problem with a quadratic objective function can be replaced with one having a linear objective function, by moving the objective function to the constraint set and by adding one more variable. (If a transformation was necessary, then our assumptions on \mathcal{R} imply that the "new" feasible region in \mathbb{R}^{n+1} , which we denote by \mathcal{R}' , is also nonempty, bounded and full dimensional. The first and third properties are trivial to establish. If \mathcal{R}' was unbounded, then from Theorems 3.1.1 and 3.1.2 it follows that \mathcal{R}' contains a ray, which then implies that \mathcal{R} contains a ray. Thus \mathcal{R}' must be bounded. Note that for more general programmes, the transformation would not necessarily preserve boundedness of the feasible region.)

Let x^* be an optimal solution to (5.1.1), and let $I(x^*)$ be the set of indices of all constraints active at x^* . Our final assumptions are that the gradients $\nabla Q_i(x^*)$, $i \in I(x^*)$, are linearly independent, that strict complementarity holds at x^* , and that

$$y^T \left(\sum_{i \in I(x^*)} \nabla^2 Q_i(x^*) \right) y > 0 \tag{5.1.2}$$

for all vectors $y \neq 0$ with $\nabla Q_i(x^*)^T y = 0$, $\forall i \in I(x^*)$. Since the objective function is linear we have that $I(x^*) \neq \emptyset$. Note that if $|I(x^*)| = n$ then (5.1.2) is satisfied vacuously as there is no $y \neq 0$ with $\nabla Q_i(x^*)^T y = 0, \forall i \in I(x^*)$. With these assumptions we can make use of Theorem 1.2.1. An immediate consequence is that x^* is the unique solution. Furthermore, if $\{x^k\}$ is generated by Algorithm A, then we can think of $\{x^k\}$ as a sequence of points on the unique isolated trajectory

$$x(r) = \operatorname{argmin}\{L(x, r)\}$$

defined on $(0, r_0]$, for some $r_0 > 0$. That is, for each $k = 1, 2, \dots$, we have $x^k = x(r_k)$ for some $r_k \in (0, r_0]$, and $x(0) = x^*$.

Definition 5.1.

We say that the sequence $\{\bar{x}^k\}$ is convergent to x^* superlinearly faster than the sequence $\{x^k\}$, if

$$\lim_{k \rightarrow \infty} \frac{\|x^* - \bar{x}^k\|}{\|x^* - x^k\|} = 0. \quad (5.1.3)$$

□

In Section 5.2 we will show how to construct the auxiliary sequence \bar{x}^k , and we will prove that it is convergent superlinearly faster to the optimal solution than the sequence $\{x^k\}$. The importance of the auxiliary sequence is that it can be used to accelerate convergence of the original sequence. In Section 5.3 we will show that the result is a superlinearly convergent method of centres for QCQP. Limited numerical experience is given in Section 5.4, and conclusions are in Section 5.5.

5.2. The Auxiliary Sequence.

Suppose that the sequence $\{x^k\}$ has been generated by Algorithm A and that

$$\lim_{k \rightarrow \infty} x^k = x^* \quad (5.2.1)$$

The auxiliary sequence $\{\bar{x}^k\}$ is defined by

$$\bar{x}^k = x^{k-1} + \lambda^{k-1}(x^k - x^{k-1}), \quad k = 1, 2, 3, \dots, \quad (5.2.2)$$

where

$$\lambda^k = \max_{\lambda > 0} \{ \lambda \mid x^k + \lambda(x^{k+1} - x^k) \in \mathcal{R} \}.$$

Since \mathcal{R} is bounded, the scalars λ^k are well-defined (see Section 5.4). We note that since both x^k and x^{k+1} are in the interior of \mathcal{R} , and since \bar{x}^{k+1} is on the boundary of \mathcal{R} it follows that $\lambda^k > 1$.

The main result, i.e., that $\{\bar{x}^k\}$ converges superlinearly faster than $\{x^k\}$ is contained in Theorem 5.2.1. The results in Lemmas 5.2.1 - 5.2.4 are required by the proof of that theorem. The lemmas are concerned with properties of the vector

$$s = \frac{x'(0)}{\|x'(0)\|}, \quad (5.2.3)$$

where

$$x'(0) = \left. \frac{dx(r)}{dr} \right|_{r=0}.$$

We note that since the functions in (5.1.1) are quadratic, it follows from Theorem 1.2.1 in Chapter 1 that $x(r)$ is continuously differentiable, on the interval $[0, r_0]$ for some $r_0 > 0$.

Lemma 5.2.1.

$$x'(0) \neq 0.$$

Proof.

Since $\nabla_x L(x(r), r) = 0$, we have

$$c - \sum_{i=1}^m u_i(r) \nabla Q_i(x(r)) = 0, \quad (5.2.4)$$

where the Lagrange multipliers $u_i(r)$ are defined by the equations

$$u_i(r) Q_i(x(r)) = r, \quad i \in I.$$

(Note that $u_i(r) < 0$ since $Q_i(x(r)) < 0$.) Under the assumptions stated in Section 5.1, both $x(r)$ and $u_i(r)$, $\forall i \in I(x^*)$, are continuously differentiable. So, we can differentiate the two previous equations with respect to r to get the matrix equation

$$\begin{bmatrix} -\sum_{i=1}^m u_i \nabla^2 Q_i(x) & -\nabla Q_1(x) & \dots & -\nabla Q_m(x) \\ u_1 \nabla^T Q_1(x) & Q_1(x) & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ u_m \nabla^T Q_m(x) & 0 & \dots & Q_m(x) \end{bmatrix} \begin{bmatrix} x' \\ u'_1 \\ \vdots \\ u'_m \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 1 \end{bmatrix}, \quad (5.2.5)$$

where all entries in (5.2.5) are functions of the penalty parameter r .

Suppose that $x'(0) = 0$. Because $x(r)$ and $u(r)$ are continuously differentiable on $[0, r_0]$ then setting $r = 0$ in (5.2.5) yields the equations

$$\sum_{i=1}^m u'_i(0) \nabla Q_i(x(0)) = 0$$

and

$$Q_i(x(0)) \cdot u'_i(0) = 1, \quad i \in I.$$

Since $u'_i(0)$ is finite for all $i \in I$ and since $Q_i(x(0)) = 0$ for all $i \in I(x^*)$, the previous equation implies that $0 = 1$, whenever $i \in I(x^*)$. (Recall that $x^* = x(0)$.) Since $I(x^*) \neq \emptyset$, this is a contradiction which establishes that $x'(0) \neq 0$. \square

Lemma 5.2.2.

$$\lim_{k \rightarrow \infty} \frac{x^k - x^*}{\|x^k - x^*\|} = s.$$

Proof.

Suppose that the hypothesis does not hold. Then there exists a vector $v \neq s$ and a set of indices k_j such that

$$\lim_{j \rightarrow \infty} \frac{x^{k_j} - x^*}{\|x^{k_j} - x^*\|} = v,$$

and such that $\|x^{k_j} - x^*\|$ is a strictly decreasing sequence. Now, for every x^{k_j} there corresponds an r_{k_j} such that $x(r_{k_j}) = x^{k_j}$. Since, by construction, $r_{k_j} \rightarrow 0$ as

$j \rightarrow \infty$, and since, by Lemma 5.2.1, $x'(0) \neq 0$, we have that

$$\begin{aligned} x'(0) &= \lim_{r \rightarrow 0^+} \frac{x(r) - x^*}{r} \\ &= \lim_{j \rightarrow \infty} \frac{x^{k_j} - x^*}{r_{k_j}} \\ &= \lim_{j \rightarrow \infty} \frac{x^{k_j} - x^*}{\|x^{k_j} - x^*\|} \cdot \frac{\|x^{k_j} - x^*\|}{r_{k_j}} \\ &= v \cdot \lim_{j \rightarrow \infty} \frac{\|x^{k_j} - x^*\|}{r_{k_j}}. \end{aligned}$$

This implies that

$$v = \frac{x'(0)}{\|x'(0)\|},$$

which contradicts $v \neq s$. \square

Lemma 5.2.3.

$$\lim_{k \rightarrow \infty} \frac{x^k - x^{k+1}}{\|x^k - x^{k+1}\|} = s.$$

Proof.

Since $x(r)$ is continuously differentiable, we can use the Mean Value Theorem to establish the existence of scalars $\theta_i^k \in (r_{k+1}, r_k) \subset [0, r]$, such that

$$\begin{aligned} &\lim_{k \rightarrow \infty} \frac{x(r_k) - x(r_{k+1})}{r_k - r_{k+1}} \\ &= \lim_{k \rightarrow \infty} \left(\frac{d x_1(\theta_1^k)}{dr}, \dots, \frac{d x_n(\theta_n^k)}{dr} \right) \\ &= \left(\frac{d x_1(0)}{dr}, \dots, \frac{d x_n(0)}{dr} \right) \\ &= x'(0). \end{aligned}$$

Since

$$\frac{x^k - x^{k+1}}{\|x^k - x^{k+1}\|} = \frac{x(r_k) - x(r_{k+1})}{\|x(r_k) - x(r_{k+1})\|} = \frac{\frac{x(r_k) - x(r_{k+1})}{r_k - r_{k+1}}}{\left\| \frac{x(r_k) - x(r_{k+1})}{r_k - r_{k+1}} \right\|},$$

it follows that

$$\lim_{k \rightarrow \infty} \frac{x^k - x^{k+1}}{\|x^k - x^{k+1}\|} = \frac{x'(0)}{\|x'(0)\|} = s.$$

\square

Lemma 5.2.4.

$$(a) \quad \nabla Q_i(x(0))^T s < 0, \quad \forall i \in I(x^*)$$

$$(b) \quad c^T s > 0.$$

Proof.

From (5.2.5) we have that

$$u_i(r) \nabla Q_i(x(r))^T x'(r) + u_i'(r) Q_i(x(r)) = 1, \quad i \in I(x^*).$$

Since $Q_i(x(0)) = 0$ for $i \in I(x^*)$ and since $u_i(r) < 0$, taking the limit of the latter equation as $r \rightarrow 0^+$ gives

$$\nabla Q_i(x(0))^T x'(0) = \frac{1}{u_i(0)} < 0.$$

This, together with (5.2.3), proves part (a). Taking the limit of (5.2.4) as $r \rightarrow 0^+$ and using strict complementary slackness gives

$$c = \sum_{i \in I(x^*)} u_i(0) \nabla Q_i(x(0))$$

Taking the inner product with $x'(0)$ and using part (a) gives

$$c^T x'(0) = \sum_{i \in I(x^*)} u_i(0) \nabla Q_i(x(0))^T x'(0) > 0.$$

This, together with (5.2.3), proves part (b). \square

Theorem 5.2.1.

$$\lim_{k \rightarrow \infty} \frac{\|x^* - \bar{x}^k\|}{\|x^* - x^k\|} = 0. \quad (5.2.6)$$

Proof.

Since \mathcal{R} is closed and bounded, the sequence $\{\bar{x}^k\}$ has at least one cluster point $z \in \mathcal{R}$. We will show that $z = x^*$ is unique. Let z be any cluster point of $\{\bar{x}^k\}$ and suppose that $z \neq x^*$. Let $J \subseteq \mathbb{N}$ be such that

$$\lim_{\substack{k \rightarrow \infty \\ k \in J}} \bar{x}^k = z.$$

Since $\lim_{k \rightarrow \infty} x^k = x^*$ we have

$$\lim_{\substack{k \rightarrow \infty \\ k \in J}} \frac{x^k - \bar{x}^k}{\|x^k - \bar{x}^k\|} = \frac{x^* - z}{\|x^* - z\|}. \quad (5.2.7)$$

Using (5.2.2), Lemma 5.2.3, and the fact that $\lambda^{k-1} > 1$, it follows that

$$\lim_{\substack{k \rightarrow \infty \\ k \in J}} \frac{x^k - \bar{x}^k}{\|x^k - \bar{x}^k\|} = \lim_{\substack{k \rightarrow \infty \\ k \in J}} \frac{x^{k-1} - x^k}{\|x^{k-1} - x^k\|} = s. \quad (5.2.8)$$

Now, (5.2.7) and (5.2.8) imply that

$$s = \frac{x^* - z}{\|x^* - z\|}.$$

Rewriting the latter equation as $z = x^* - \|x^* - z\|s$, and using the fact that the $Q_i(x)$ are convex, we have

$$Q_i(z) \geq Q_i(x^*) - \|x^* - z\| \nabla Q_i(x^*)^T s,$$

which, together with Lemma 5.2.4(a), gives $Q_i(z) > 0$. This contradicts $z \in \mathcal{R}$ and so establishes that $z = x^*$. Thus,

$$\lim_{k \rightarrow \infty} \bar{x}^k = x^*. \quad (5.2.9)$$

Let $M = \{k \mid \bar{x}^k = x^*\}$ and let v and $K \subseteq \mathbb{N} \setminus M$ be such that

$$\lim_{\substack{k \rightarrow \infty \\ k \in K}} \frac{x^* - \bar{x}^k}{\|x^* - \bar{x}^k\|} = v. \quad (5.2.10)$$

From (5.2.9) we conclude that there exists an infinite set $L \subseteq K$ and an index $i \in I(x^*)$ such that $Q_i(\bar{x}^k) = 0$ for all $k \in L$. It now follows from Taylor's theorem that

$$\nabla Q_i(x^*)^T \frac{(\bar{x}^k - x^*)}{\|\bar{x}^k - x^*\|} = - \frac{(\bar{x}^k - x^*)^T \nabla^2 Q_i(\eta)(\bar{x}^k - x^*)}{2 \|\bar{x}^k - x^*\|},$$

for some $\eta \in (\bar{x}^k, x^*)$. Since $Q_i(x)$ is twice continuously differentiable, we can take the limit as $k \rightarrow \infty$, $k \in L$, to get

$$\nabla Q_i(x^*)^T v = 0.$$

From Lemma 5.2.4(a) we can conclude that $v \neq \pm s$.

We can now establish (5.2.6). If $\mathbb{N} \setminus M$ is finite, then (5.2.6) holds. Now suppose that $\mathbb{N} \setminus M$ is infinite. For $k \in \mathbb{N} \setminus M$ the points x^k , \bar{x}^k , and x^* are not collinear, and so they define a triangle. Applying the Law of Sines gives

$$\frac{\|x^* - \bar{x}^k\|^2}{\|x^* - x^k\|^2} = \frac{1 - \left(\left(\frac{x^k - \bar{x}^k}{\|x^k - \bar{x}^k\|} \right)^T \left(\frac{x^k - x^*}{\|x^k - x^*\|} \right) \right)^2}{1 - \left(\left(\frac{x^* - \bar{x}^k}{\|x^* - \bar{x}^k\|} \right)^T \left(\frac{x^k - \bar{x}^k}{\|x^k - \bar{x}^k\|} \right) \right)^2}$$

Now, from $v \neq \pm s$, (5.2.10), and Lemma 5.2.2, it follows that

$$\begin{aligned} \lim_{\substack{k \rightarrow \infty \\ k \in K}} \frac{\|x^* - \bar{x}^k\|^2}{\|x^* - x^k\|^2} &= \lim_{\substack{k \rightarrow \infty \\ k \in K}} \frac{1 - \left(\left(\frac{x^k - \bar{x}^k}{\|x^k - \bar{x}^k\|} \right)^T \left(\frac{x^k - x^*}{\|x^k - x^*\|} \right) \right)^2}{1 - \left(\left(\frac{x^* - \bar{x}^k}{\|x^* - \bar{x}^k\|} \right)^T \left(\frac{x^k - \bar{x}^k}{\|x^k - \bar{x}^k\|} \right) \right)^2} \\ &= \frac{1 - (s^T s)^2}{1 - (v^T s)^2} = 0. \end{aligned}$$

Because the latter equality holds for any cluster point v , we have that

$$\limsup_{k \rightarrow \infty} \frac{\|x^* - \bar{x}^k\|}{\|x^* - x^k\|} = 0.$$

Since the sequence in (5.2.6) has only positive terms for $k \notin M$, then the result follows. \square

5.3. Superlinearly Convergent Algorithm.

The elements in the sequences $\{\bar{x}^k\}$ and $\{x^k\}$ only depend upon previous elements in the original sequence $\{x^k\}$, and not on the elements in the auxiliary sequence $\{\bar{x}^k\}$. In the next theorem we show how the elements in the auxiliary sequence can be used to construct a sequence of points on the trajectory which is at least superlinearly convergent.

Let us define the sequence

$$\tilde{x}^k = x^k + \alpha_k(\bar{x}^k - x^k) \quad (5.3.1a)$$

where $0 < \alpha_k < 1$, $\alpha_k < \alpha_{k+1}$ and

$$\lim_{k \rightarrow \infty} \alpha_k = 1. \quad (5.3.1b)$$

We will use the sequence $\{\tilde{x}^k\}$ to define the truncation levels and the truncated feasible regions

$$\mathcal{R}_k = \{x \in \mathcal{R} | c^T x \leq c^T \tilde{x}^k\}.$$

in the subsequent iterations.

Lemma 5.3.1.

$$\lim_{k \rightarrow \infty} \frac{\|\tilde{x}^k - x^*\|}{\|x^k - x^*\|} = 0.$$

Proof.

By Theorem 5.2.1

$$\lim_{k \rightarrow \infty} \frac{\|\bar{x}^k - x^*\|}{\|x^k - x^*\|} = 0. \quad (5.3.2)$$

Consider the triangle with vertices x^* , \bar{x}^k and x^k . Using the triangle inequality together with (5.3.2) gives

$$\lim_{k \rightarrow \infty} \frac{\|\bar{x}^k - x^k\|}{\|x^k - x^*\|} = 1. \quad (5.3.3)$$

By the formulas (5.3.1) we get

$$\|\tilde{x}^k - x^k\| = \alpha_k \|\bar{x}^k - x^k\|.$$

The latter along with (5.3.3) and (5.3.1a) implies

$$\lim_{k \rightarrow \infty} \frac{\|\tilde{x}^k - x^k\|}{\|x^k - x^*\|} = 1$$

and hence

$$\lim_{k \rightarrow \infty} \frac{\|\tilde{x}^k - \bar{x}^k\|}{\|x^k - x^*\|} = 0. \quad (5.3.4)$$

Now (5.3.2), (5.3.4), and the triangle inequality applied to the triangle with vertices x^* , \tilde{x}^k and \bar{x}^k gives

$$\lim_{k \rightarrow \infty} \frac{\|x^* - \tilde{x}^k\|}{\|x^k - x^*\|} = 0.$$

□

Theorem 5.3.1.

If the sequence $\{x^k\}$ is such that $c^T x^{k+1} \leq c^T \tilde{x}^k$, and that $\lim_{k \rightarrow \infty} x^k = x^*$, then

$$\lim_{k \rightarrow \infty} \frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|} = 0.$$

Proof.

Let θ_k be the angle between the vectors c and $(x^{k+1} - x^*)$ and let $\bar{\theta}_k$ be the angle between the vectors c and $(\tilde{x}^k - x^*)$. Note that since x^{k+1} is not an optimal solution it follows that $\theta_k \neq 0$. We have that

$$\begin{aligned} \frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|} &= \frac{\frac{c^T (x^{k+1} - x^*)}{\cos \theta_k \|c\|}}{\|x^k - x^*\|} \\ &\leq \frac{\frac{c^T (\tilde{x}^k - x^*)}{\cos \bar{\theta}_k \|c\|}}{\|x^k - x^*\|} \\ &= \frac{\|\tilde{x}^k - x^*\| (\cos \bar{\theta}_k / \cos \theta_k)}{\|x^k - x^*\|}. \end{aligned}$$

Using Lemma 5.3.1 and Lemma 5.2.4(b), which establishes that there exists an $\epsilon > 0$ and a $K > 0$ such that $\cos \theta_k > \epsilon$ for $k > K$, we get

$$\begin{aligned} \lim_{k \rightarrow \infty} \frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|} &= \lim_{k \rightarrow \infty} \frac{\|\tilde{x}^k - x^*\| (\cos \bar{\theta}_k / \cos \theta_k)}{\|x^k - x^*\|} \\ &= 0 \cdot \lim_{k \rightarrow \infty} (\cos \bar{\theta}_k / \cos \theta_k) \\ &= 0. \end{aligned}$$

In fact the

$$\lim_{k \rightarrow \infty} (\cos \bar{\theta}_k / \cos \theta_k)$$

may not exist but the sequence is bounded. \square

For the method of centers, it is easy to choose the sequence $\{x^k\}$ satisfying the conditions of Theorem 5.3.1. We simply change the truncated feasible region (5.1.3) to

$$\mathcal{R}(z_k) = \{x \in \mathcal{R} \mid c^T x \leq c^T \bar{x}^k\},$$

where $z_k = c^T \bar{x}^k$. In order to determine an analytic center of the region \mathcal{R}_k the point \hat{x}_k can be used as an initial approximation. While it is easy to adjust the method of analytic centers to achieve a superlinear rate of convergence, it is a different story for the barrier function method. The difficulty is in determining the sequence of penalty parameters $\{\bar{r}_k\}$ to guarantee $c^T x^{k+1} \leq c^T \bar{x}^k$. We first note that the parameter r^{k+1} given in (5.1.4) is such that $x^{k+1} = x(r^{k+1})$ satisfies $c^T x^{k+1} \leq c^T x^k$. So, to get $x^{k+1} = x(\bar{r}^{k+1})$ with $c^T x^{k+1} \leq c^T \bar{x}^k$, we can choose

$$\bar{r}^{k+1} = \frac{c^T (\bar{x}^k - x^{k+1})}{m}. \quad (5.3.5)$$

Of course, the difficulty here is that \bar{r}^{k+1} is a function of $x^{k+1} = x(\bar{r}^{k+1})$. Therefore, an implementation requires a method to approximate the solution \bar{r}^{k+1} to (5.3.5). The superlinearly convergent algorithm suggested by this section is given below.

Algorithm E: A Superlinearly Convergent Modified Method of Centers

$x^0 \in \text{int}(\mathcal{R})$, $k = 0$

repeat

set $\tilde{x}^k = x^k + \alpha_k(\bar{x}^k - x^k)$, $z_k = c^T \tilde{x}^k$
Find $x^{k+1} = \text{argmin} \{ F(x, z_k) \mid x \in \text{int}(\mathcal{R}(z_k)) \}$
if convergence criteria are not satisfied **then**
 $k := k + 1$
else stop := true

until stop $\{x^* \approx x_k\}$

\square

When solving for x^{k+1} , we can use $\hat{x}^k = x^k + \frac{1}{2}(1 + \alpha_k)(\bar{x}^k - x^k)$ as an initial approximation to x^k . Also, we can use $\alpha_k = \frac{k}{k+1}$.

We note that the Algorithm E can be classified as a large-step method of centers. Another method in this class has been proposed by Den Hertog et al. in [12]. Both methods use several Newton iterations to determine the iterate x^{k+1} . In [12] the stopping criteria for Newton's method is such that the algorithm has polynomial time complexity. In our algorithm, we need a more precise estimate for x^{k+1} . However, the extra precision allows us to construct a point on the auxiliary sequence, from which we calculate \bar{x}^k . Thus, our algorithm makes deeper cuts and achieves a superlinear rate of convergence.

5.4. Computational Results.

To see the potential of Algorithm E, we compare it's performance on three test problems to that of the method of centers as given in Algorithm B.

In order to calculate the auxiliary sequence $\{\bar{x}^k\}$ we need explicit expressions for the scalars λ^k in (5.2.2). Suppose that x^k and x^{k+1} are given. Now define $\delta^k = x^k - x^{k+1}$, and $M_k = \{i \in I | B_i \delta^k = 0\}$. Now, for each $i \in M_k$ we define

$$\lambda_i^k = \frac{b_i - a_i^T x^k}{a_i^T \delta^k}$$

and for each $i \notin M_k$ we define

$$\lambda_i^k = \frac{-\nabla Q_i(x^k)^T \delta^k + \sqrt{[\nabla Q_i(x^k)^T \delta^k]^2 - 2Q_i(x^k)(\delta^k)^T B_i \delta^k}}{(\delta^k)^T B_i \delta^k}.$$

We then set

$$\lambda^k = \min_{\substack{i \in I \\ \lambda_i^k > 0}} \lambda_i^k.$$

From the above, we see that the calculation of \bar{x}^{k+1} requires $O(n^2 m)$ operations. This is an order of magnitude less than the cost of computing the element x^k . For example, the cost of a single Newton step is $O(n^3 + mn^2)$.

As our convergence criteria we use $\|x^{k+1} - x^k\| \leq \epsilon$, or $|c^T x^{k+1} - c^T x^k| \leq \epsilon$, where $\epsilon = 10^{-9}$. The examples are given below. Note that all constraint functions are quadratic.

Example 5.1 [18]. We have that $m = 5$, $n = 6$,

$$B_1 = \begin{bmatrix} I_5 & 0 \\ 0 & 0 \end{bmatrix},$$

where I_5 is an unit matrix,

$$B_2 = \begin{bmatrix} 9 & 0 & 6 & 3 & 3 & 0 \\ 0 & 16 & -4 & -16 & -16 & 0 \\ 6 & -4 & 5 & 6 & 6 & 0 \\ 3 & -16 & 6 & 33 & -3 & 0 \\ 3 & -16 & 6 & -3 & 78 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$B_3 = \begin{bmatrix} 1 & -10 & 3 & 4 & 5 & 0 \\ -10 & 136 & -30 & -94 & -26 & 0 \\ 3 & -30 & 18 & 18 & 15 & 0 \\ 4 & -94 & 18 & 122 & -15 & 0 \\ 5 & -26 & 15 & -15 & 58 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$B_4 = B_5 = [0]$, $c^T = (0, 0, 0, 0, 0, 1)$, $a_1^T = (-1, -1, -1, -1, -1, -1)$, $a_2^T = (2, 2, 1, 0, -2, 0)$, $a_3^T = (0, 0, 0, -2, -2, 0)$, $a_4^T = (-1, 0, 0, 0, 0, 0)$, $a_5^T = (0, -1, 0, 0, 0, 0)$, $b_1 = -2$, $b_2 = -1$, $b_3 = -4$, $b_4 = b_5 = 0$, and $x^{*T} = (0.00000, 0.33716, 0.10883, 0.21507, 0.12293, -2.69049)$.

Example 5.2 [10]. We have $m = 3$, $n = 3$,

$$B_1 = \begin{bmatrix} 5 & 7 & 0 \\ 7 & 13 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$B_2 = \begin{bmatrix} 5 & -1 & 0 \\ -1 & 10 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$B_3 = \begin{bmatrix} 4 & -2 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$c^T = (0, 0, 1)$, $a_1^T = (18, -32, -1)$, $a_2^T = (2, 3, 0)$, $a_3^T = (-2, 1, 0)$, $b_1 = 0$, $b_2 = -11.5$, $b_3 = -1$, and $x^{*T} = (1, 1, -34)$. \square

Example 5.3 [23]. We have $m = 1$, $n = 19$,

$B_1 = \text{diag}(1, 2, 3, 4, 5, 6, 7, 8, 9, 10, .9, .8, .7, .6, .5, .4, .3, .2, 0)$,

$a_1^T = -e_{19}$, $b_1 = 0$, $c^T = e_{19}$, and $x^* = 0$, where $e_{19} \in \mathbb{R}^{19}$ is a vector with the 19-th coordinate equal to 1 and remaining coordinates equal to zero. \square

Computational results are reported in Tables 1 and 2. The entries in Table 1 indicate the number of iterations and the total number of Newton steps required to obtain an optimal solution exact with 10 decimal digits. The entries in Table 2 indicate the values of $\frac{c^T(x^k - x^*)}{c^T(x_0 - x^*)}$, which are relative distances of x^k from the optimal solution x^* , expressed in terms of the objective function value.

Example	1	2	3
Algorithm B	40 (188)	48 (259)	51 (224)
Algorithm E	13(31)	14 (60)	4 (4)

Table 1: Number of iterations and Newton steps

The results are very encouraging for Algorithm E. It requires both fewer iterations and fewer Newton steps than Algorithm B.

		Example 1		Example 2		Example 3	
		Algorithm		Algorithm		Algorithm	
		B	E	B	E	B	E
	1	3E-01	3E-01	4E-01	4E-01	7E-01	7E-01
I	2	2E-01	7E-02	2E-01	1E-01	4E-01	3E-01
t	3	8E-02	2E-02	1E-01	3E-02	3E-01	6E-02
e	4	4E-02	4E-03	7E-02	6E-03	2E-01	1E-02
r	5	2E-02	8E-05	4E-02	9E-04	1E-01	2E-03
a	6	1E-02	9E-06	2E-02	1E-04	7E-02	3E-04
t	7	6E-03	7E-07	1E-02	1E-05	5E-02	3E-05
i	8	3E-03	5E-08	7E-03	1E-06	3E-02	4E-06
o	9	2E-03	7E-09	4E-03	8E-08	2E-02	4E-07
n	10	9E-04	7E-09	2E-03	6E-09	1E-02	3E-08
	11	5E-04	4E-09	1E-03	4E-10	8E-03	3E-09
	12	3E-04	4E-09	6E-04	2E-11	5E-03	2E-10

Table 2: The values of $\frac{c^T(x^k - x^*)}{c^T(x_0 - x^*)}$

The results in Table 2 certainly support the claim of superlinear convergence. In all three examples, Algorithm E takes no more than 4 iterations for an improvement from 10^{-5} to 10^{-9} . Also, for iterations beyond the fifth, the numbers in the columns for Algorithm E are at least 10 times smaller than those for Algorithm B.

Of course, no conclusions can be made on the basis of these tests. However, they do indicate that further study is warranted.

5.5. Summary.

We presented a technique for using the sequence of points on the optimal trajectory to create an auxiliary sequence on the boundary of the feasible region which converges superlinearly faster. We showed that this auxiliary sequence can be con-

structed for the method of centers as well as for the logarithmic barrier function method. The cost of calculating the auxiliary sequence is $O(n^2m)$ so that the polynomial time property of the original algorithm is maintained. In addition, we propose a superlinearly convergent algorithm which stems from using the auxiliary sequence to increase the rate of convergence of the original sequence $\{x^k\}$. We also provide some numerical results for the superlinearly convergent Algorithm E proposed in Section 5.3.

CHAPTER 6

CONCLUDING REMARKS

6.1. Introduction.

This chapter serves both as a summary of the contribution of this thesis and as a guide to future research.

6.2. Contribution of Thesis.

This thesis provides new results on the minimal representation of quadratically constrained convex feasible regions. We introduced the definition of minimal representation and we proved that a representation is minimal if and only if it contains no redundant constraints, no implicit equalities and no pseudo-quadratic constraints. We also outlined the steps of a procedure that could be used to determine a minimal representation. We indicated that there are potential benefits in the determination of a minimal representation before the application of interior point methods for the solution of QCQP problems. As Telgen's results proved to be a major building block for the facet generation algorithms in combinatorial optimization, these results may be of significance in furthering non-linear programming along the same lines. In addition, the benefits of one of the results on the faces of \mathcal{R} go beyond the proof of the minimal representation theorem. Theorem 2.4.1 allows an extension of the HD method, which detects necessary constraints in linear systems, to systems of inequalities with convex analytic functions.

This thesis presents a new algorithm [see also 8], which can be used to determine whether or not a convex quadratic function is bounded from below over a feasible region defined by convex quadratic constraints. If there are m constraints and n variables, the algorithm terminates after at most $\min\{m, n\}$ iterations. In addition, the algorithm has the advantage of providing a mechanism to reduce both the number of constraints and the dimension of the problem.

This thesis presents a new method for dealing with the application of a method

of centers to QCQP problems having a faulty feasible regions, that is regions that are not full dimensional or that are unbounded.

This thesis provides a technique for improving the rate of convergence of interior point methods applied to QCQP problems. The approach can also be extended to convex programs with smooth constraints. It uses the sequence of points on the optimal trajectory to create the auxiliary sequence on the boundary of the feasible region, which converges superlinearly faster. The auxiliary sequence is then used to create a new sequence with a superlinear rate of convergence. We also provide some numerical results to demonstrate the acceleration technique.

6.3. Future Research Directions.

While our interest in the problems addressed in the thesis was motivated by the potential for computational savings in the solution of the QCQP problem, this thesis is mainly a theoretical contribution. Further research will be aimed at computational testing. However, this will not be our only goal. We have already obtained extensions of the results to convex constraints under differentiability assumptions.

A generalization of the minimal representation results to convex analytic functions has already been done, and a paper on this topic is in preparation [43]. Although it can be shown that similar results for arbitrary convex functions do not hold, we will consider future research on nonconvex regions defined by analytic functions.

Extensions of the unboundedness results have already been completed for convex analytic functions [38]. However, some implementation problems remain. This problem is linked to the problem of finding a cone of recession of a convex function, and, also, of a convex region. This area seems to be promising, and there are no techniques in the existing literature to handle this problem.

Furthermore, we plan to work on the infeasibility problem for quadratically constrained convex regions. We have already obtained some results on this problem.

Our results make use of implicit equalities in a corresponding linear system. To the best of our knowledge, except for traditional approaches [19], this problem has not been yet considered in the existing literature. The traditional approaches rely on the solution of a nonlinear problem, with a structure similar to that of the original problem. It was shown in [45] that the feasibility problem for a quadratically constrained convex region is equivalent to the feasibility problem for related semidefinite programming problems.

This indicates that another possibility is to extend our results to semidefinite programming problems. Our expectations are based on the fact that any convex quadratically constrained problem can be represented as a semidefinite programming problem. It has been already shown [45, 54] that the theory of semidefinite programming closely parallels linear programming theory and that many algorithms for solving LP have generalizations that handle semidefinite programming problems.

FIGURES

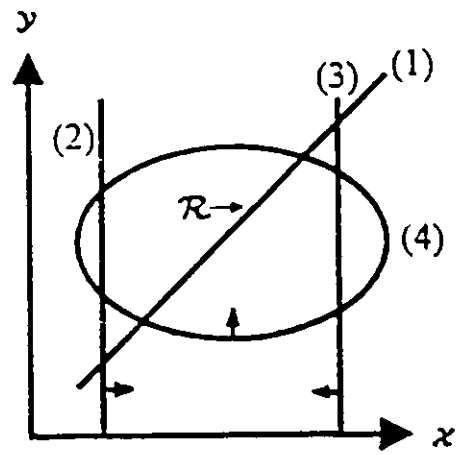


Figure 2.1(a): Constraint (4) can be replaced with linear constraints.

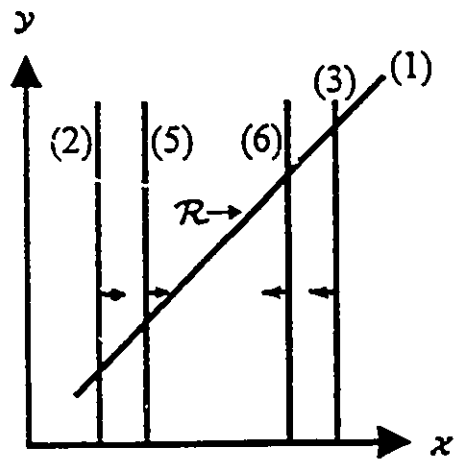


Figure 2.1(b): Constraints (2) and (3) are redundant.

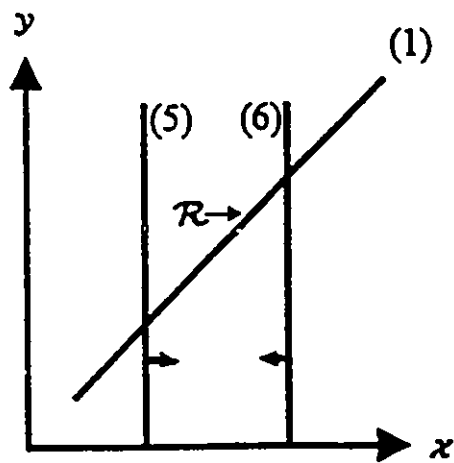


Figure 2.1(c): A minimal representation of \mathcal{R} .

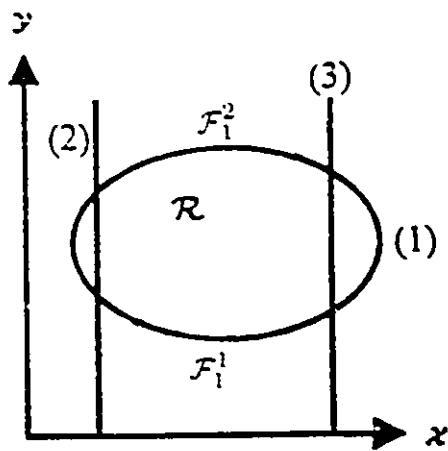


Figure 2.1(d): A face, \mathcal{F}_1 , having more than one component.

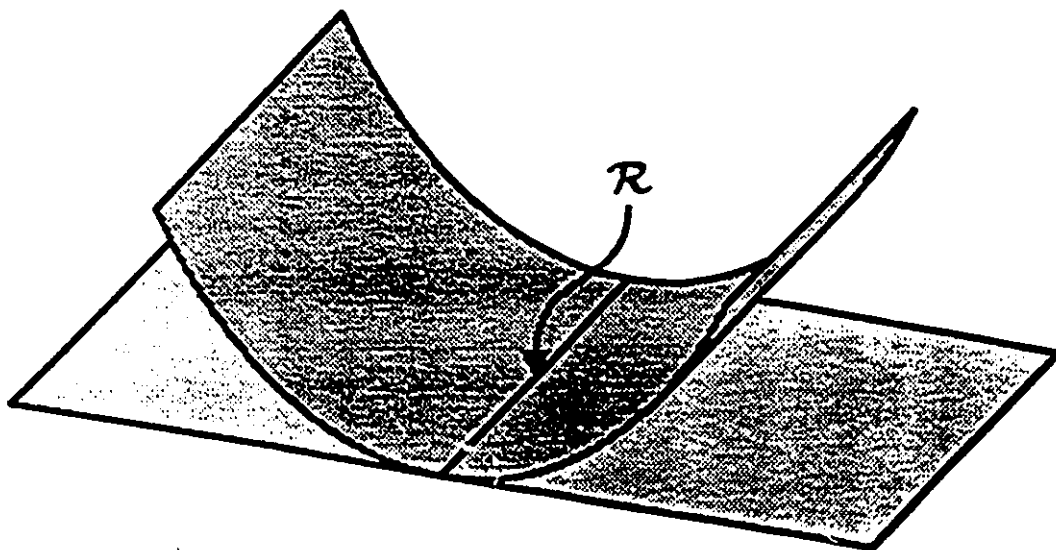


Figure 2.2: The quadratic constraint can be replaced with one linear equality.

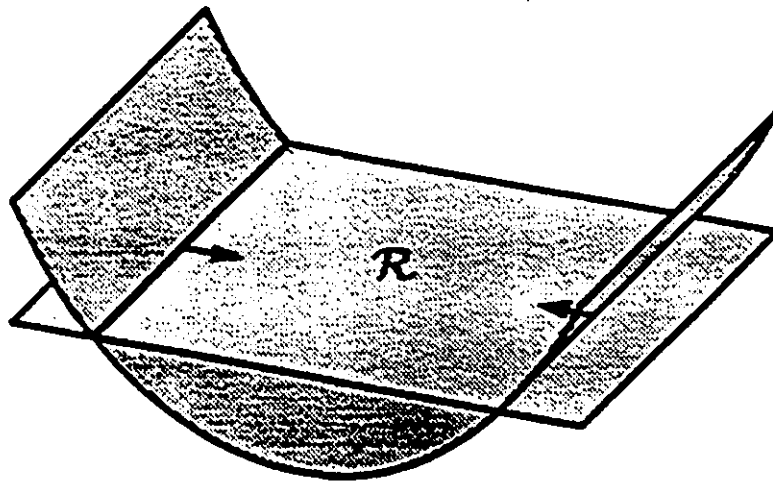


Figure 2.3: The quadratic constraint can be replaced with two linear inequality constraints.

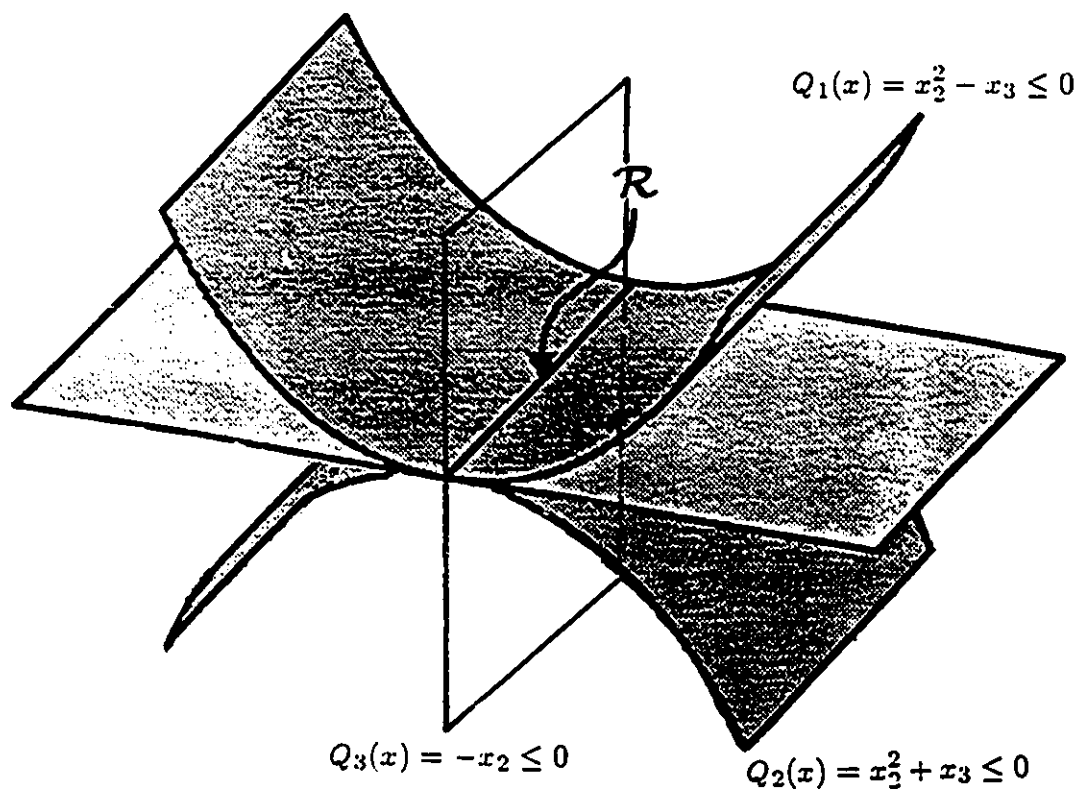


FIGURE 2.4: Illustration for Algorithm C.

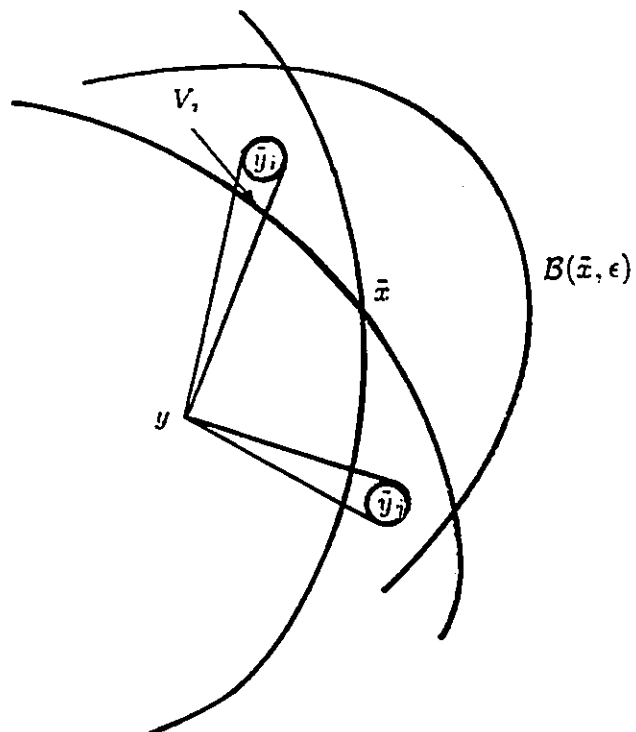


Figure 2.5 Illustration for the proof of Theorem 2.4.2

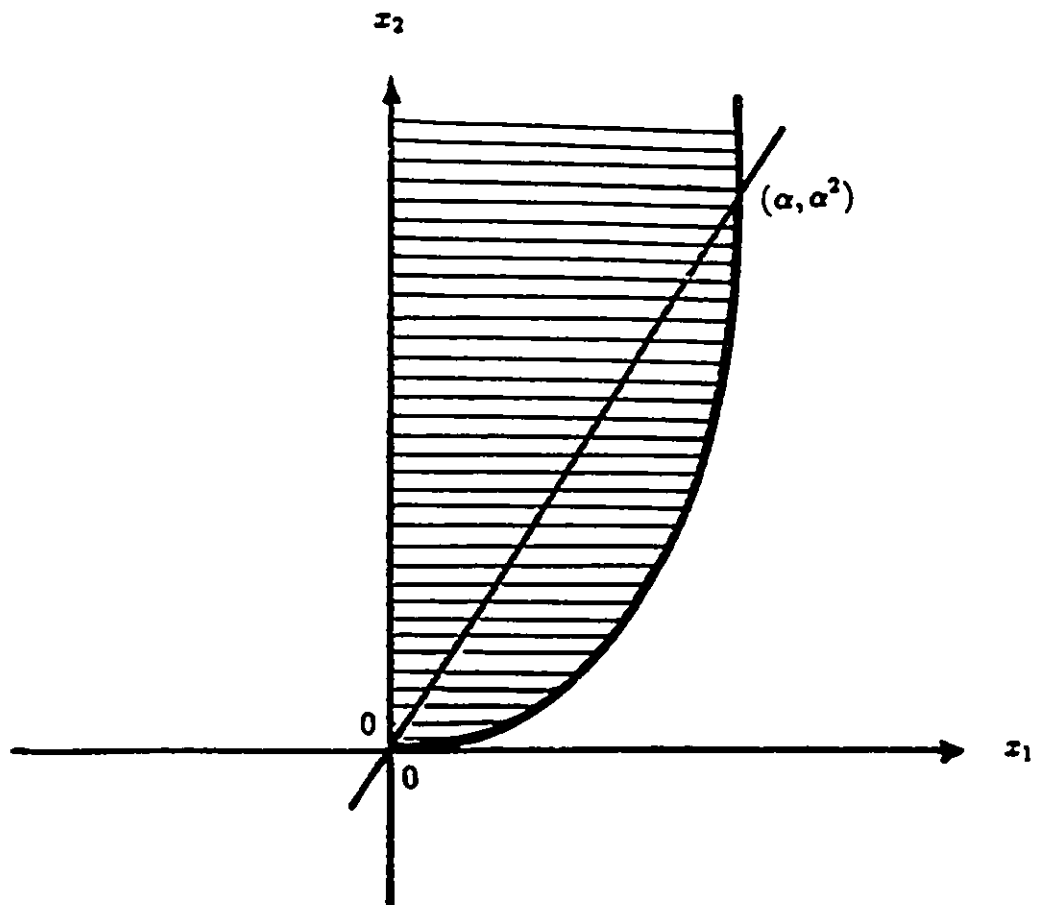


Figure 3.1: Illustration for Example 3.1.

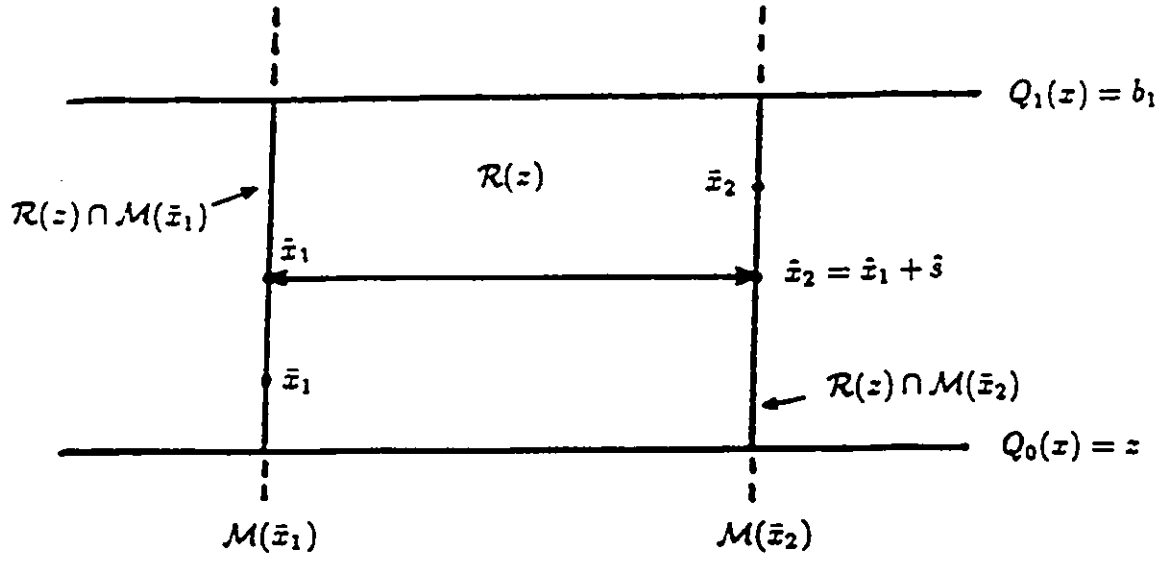


Figure 4.1: $\mathcal{R}(z)$ has an infinity of analytic centres.

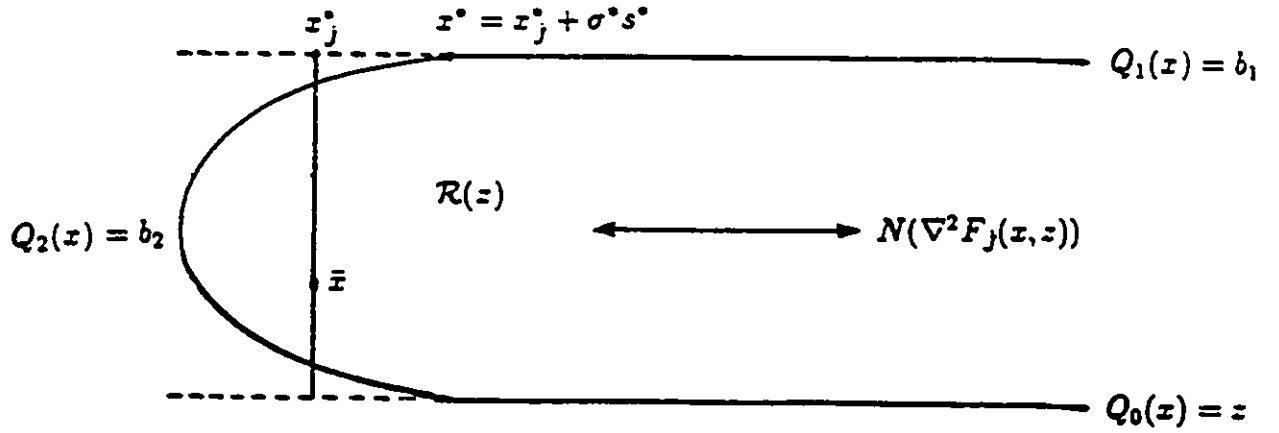


Figure 4.2: $\mathcal{R}(z)$ has no analytic centre but the QCQP has a solution.

REFERENCES

- [1] E. R. Barnes, "A variation on Karamarkar's algorithm for solving linear programming problems," *Mathematical Programming* 36 (1986) 174-182.
- [2] D.P. Baron, *Quadratic programming with quadratic constraints*, Naval Research Logistics Quarterly 19 (1972).
- [3] A. Boneh and A. Golan, *Constraints redundancy and feasible region boundedness by random feasible point generator (RFPG)*, Contributed paper, Third European Congress on Operations Research, Euro III (Amsterdam, The Netherlands, April 1979).
- [4] P.M. Camerini, S.J. Chung, K.G. Murty, *On checking of unboundedness of functions*, The Arabian Journal of Science and Engineering, Vol. 16, No., 2B(1991), 255-262.
- [5] R.J. Caron and J.F. McDonald, *A new approach to the analysis of random methods for detecting necessary linear inequality constraints*, Mathematical Programming, 43(1989) 97-102.
- [6] R.J. Caron, J.F. McDonald and C.M. Ponik, *A degenerate extreme point strategy for the classification of linear constraints as redundant or necessary*, Journal of Optimization Theory and Applications, Vol. 62, No.2, 1989.
- [7] R.J. Caron and W.T. Obuchowska, *Unboundedness of a convex quadratic function subject to concave and convex quadratic constraints*, European Journal of Operations Research, (1992), pp. 114-123.
- [8] R.J. Caron and W.T. Obuchowska, *An algorithm to determine boundedness of quadratically constrained convex quadratic programmes*, Windsor Mathematics and Statistics Report WMSR 92-15, University of Windsor, Windsor, 1992, to appear in the European Journal of Operational Research, Vol. 80, (1994).
- [9] R.J. Caron and W.T. Obuchowska, *Quadratically constrained convex quadratic*

- programmes: Faulty feasible regions*, WMSR, 92-05 , 1992.
- [10] F. Cole, J.G. Ecker, and W. Gochet, *A reduced gradient method for quadratic programming with quadratic constraints and lp-approximation problems*, European Journal of Operations Research 9 (1982) 194-203.
 - [11] D. Den Hertog, C. Roos, and T. Terlaky, *A potential reduction variant of Renegar's short-step path-following method for linear programming*, Linear Algebra and Its Applications, 152 (1991), pp. 43-68.
 - [12] D. Den Hertog, C. Roos, and T. Terlaky, *A large-step analytic center method for a class of smooth convex programming problems*, SIAM Journal on Optimization, 2 (1992), pp. 55-70.
 - [13] D. Den Hertog, C. Roos, and T. Terlaky, *On the classical logarithmic barrier function method for a class of smooth convex programming problems*, Journal of Optimization Theory and Applications, 73 (1992), pp. 1-25.
 - [14] I. I. Dikin, "Iterative solution of problems of linear and quadratic programming," Soviet Mathematics Doklady 18 (1967) 674-675.
 - [15] J.J. Dongarra, C.B. Moler, J.R. Bunch, and G.W. Stewart, Linpack: Users' Guide, (Siam, Philadelphia, 1979).
 - [16] J.G. Ecker and R.D. Niemi, *A dual method for quadratic programs with quadratic constraints*, SIAM Journal of Applied Mathematics 28 (1975) 568-576.
 - [17] U. Eckhardt, *Representation of convex sets*, Proceedings of International Symposium on External Methods and Systems Analysis, Austin, Texas, 1979.
 - [18] S.C. Fang and J.R. Rajasekera, *Controlled perturbations for quadratically constrained quadratic programs*, Mathematical Programming 36 (1986) 276-289.
 - [19] A.V. Fiacco and G.P. McCormick, *Nonlinear Programming: a Sequential Unconstrained Minimization Technique*, John Wiley and Sons, New York, 1968.
 - [20] P.E. Gill, W. Murray, M.A. Saunders, J.A. Tomlin, and M.H. Wright, *On projected Newton-Barrier methods for linear programming and on equivalence to*

- Karmarkar's projective method*, Mathematical Programming, 36 (1986), pp. 183-209.
- [21] D. Goldfarb, S. Liu, and S. Wang, *A logarithmic barrier function algorithm for quadratically constrained convex quadratic programming*, SIAM Journal on Optimization, 1 (1991), pp. 252-267.
 - [22] C.C. Gonzaga, *Large-step path-following methods for linear programming. Part I. Barrier Function Method*, SIAM Journal on Optimization, 1 (1991), pp. 268-279.
 - [23] D.M. Himmelblau, "Applied Nonlinear Programming", McGraw-Hill Book Company, 1972.
 - [24] W. Hock, K. Schittkowski, *Test examples for nonlinear programming codes*, Springer-Verlag, 1981.
 - [25] P. Huard, *Resolution of mathematical programming with nonlinear constraints by the method of centers*, in Nonlinear Programming, J. Abadie, ed., Amsterdam, 1967, North Holland, pp. 207-219.
 - [26] F. Jarre, *On the convergence of the method of analytic centers when applied to convex quadratic programs*, Mathematical Programming, 49 (1991), pp. 341-358.
 - [27] F. Jarre, *The Method of Analytic Centers for Smooth Convex Programs*, Ph.D. Thesis, Institut für Angewandte Mathematik und Statistik, Universität, Würzburg, Germany, 1989.
 - [28] N. Karmarkar, *A new polynomial time algorithm for linear programming*, Combinatorica, 4 (1984), pp. 373-395.
 - [29] V. V. Kovacevic-Vujcic, "Improving the rate of convergence of interior point methods for linear programming," *Mathematical Programming* 52 (1991) 467-479.
 - [30] E. Kranich, "Interior point methods for Mathematical Programming: A bibliography," Diskussionsbeitrag No. 171, University of Wuppertal (Wuppertal, Germany, 1992).

- [31] F.A. Lootsma, *Constrained optimization via parameter-free penalty functions*, Phillips Research Report 23 (1968) 424 - 437.
- [32] I.J. Lustig, *Feasibility issues in primal-dual interior point method for linear programming*, Mathematical Programming, 49 (1990), pp. 145-162.
- [33] O.L. Mangasarian, *Nonlinear Programming* (McGraw Hill, New York, 1969).
- [34] S. Mehrotra and J. Sun, "A method of centers for quadratically constrained convex quadratic programmes", Technical Report 88-01, Department of Industrial Engineering and Management Sciences Northwestern University (Evanston, IL, 1988).
- [35] S. Mehrotra and Y. Sun, *An algorithm for convex quadratic programming that requires $O(n^{3.5}L)$ arithmetic operations*, Mathematics of Operations Research, 15 (1990), pp. 352-363.
- [36] S. Mehrotra and Y. Sun, *On computing the center of a convex quadratically constrained quadratic set*, Mathematical Programming, 50 (1991) 81-83.
- [37] K.G. Murty, *Linear Complementarity, Linear and Nonlinear Programming*, Helderman, Verlag, Berlin, 1988.
- [38] K. G. Murty, W. T. Obuchowska, *Problem of unboundedness of convex functions*, presented at the International Symposium on Mathematical Programming, Ann Arbor, Michigan, USA, August 15-19, 1994.
- [39] Y.E. Nesterov and A.S. Nemirovsky, *Interior-Point Polynomial Algorithms in Convex Programming*, SIAM, 1994.
- [40] W. Obuchowska, "Badanie ograniczonosci funkcji celu w zagadnieniu rozszerzonej optymalizacji kwadratowej", Research Papers of the Academy of Economics (Wroclaw, Poland, 1985)
- [41] W.T. Obuchowska, R.J. Caron, *Improving the rate of convergence of interior point methods for smooth convex programmes*, WMSR No. 92-11, July 1992.
- [42] W.T. Obuchowska and R.J. Caron, *Minimal representation of quadratically con-*

- strained convex feasible regions*, to appear in Mathematical Programming.
- [43] W.T. Obuchowska, *Minimal representation of convex regions* presented at the International Symposium on Mathematical Programming, Ann Arbor, Michigan, USA, August 15-19, 1994.
 - [44] E. Phan-huy-Hao, *Quadratically constrained quadratic programming: some applications and a method for solution*, Zeitschrift für Operations Research 8 (1982) 105-119.
 - [45] M.V. Ramana, *An algorithmic analysis of multiquadratic and semidefinite programming problems*, Ph.D. thesis, The Johns Hopkins University, October 1993.
 - [46] J. Renegar, *A polynomial-time algorithm based on Newton's method for linear programming*, Mathematical Programming 40 (1988), pp 59-93.
 - [47] C. Roos and J.P. Vial, *Long steps with the logarithmic penalty barrier function in linear programming*, in: J. Gabszewicz, J.F. Richard, and L. Wolsey, eds., *Economic Decision Making: Games, Economics and Optimization*, Elsevier Science Publisher, Amsterdam, (1990), pp. 433-441.
 - [48] R.T. Rockafellar, *Convex Analysis* (Princeton University Press, New Jersey, 1970).
 - [49] S. Sonnevend, *An analytic center for polyhedrons and new classes of global algorithms for linear (smooth convex) programming* Lecture notes in Control and Information Sciences 84, Springer-Verlag (1985), pp.866-876.
 - [50] J. Telgen, *Minimal representation of convex polyhedral sets*, Journal of Optimization Theory and Applications 38(1982), 1-24.
 - [51] J. Telgen, *Identifying redundant constraints and implicit equalities in systems of linear constraints*, Management Science 29 (1983) 1209-1222.
 - [52] H. Wolkowicz, *Method of reduction in convex programming*, Journal of optimization Theory and Applications, Vol. 40, No 3, July 1983.
 - [53] P.M. Vaidya, *An algorithm for linear programming which requires $((m+n)n^2 +$*

- $(m + n)^{1.5n}L$) *arithmetic operations*, Mathematical Programming. 47 (1990), pp. 175-201.
- [54] Vanderberghe L., Boyd S., *Positive-Definite Programming*, in Mathematical Programming - State of the Art, ed. by Birge J. R., and Murty K. G., 1994.
- [55] C.R. Vogel, *A constrained least squares method for nonlinear ill-posed problems*, SIAM Journal on Control and Optimization 28(1990) 34-49.

VITA AUCTORIS

Name	Wiesława T. Obuchowska
1970	Completed Master's degree in Mathematics (Numerical Methods) from Wrocław University, Wrocław, Poland
1980	Completed Ph.D. in Mathematical Economics from the Academy of Economics, Wrocław, Poland
1980-89	Associate Professor of Economics, Academy of Economics Wrocław, Poland
1989	M.Sc. in Applied Mathematics (Optimization), University of Windsor, Windsor, Ontario
1995	Ph.D. in Applied Mathematics (Optimization), University of Windsor, Windsor, Ontario